

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**



FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGCCTGTCAGTTGACCTACCCCTTGACACCTACCCTAA
GGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTGTGGATG
ATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCATATTCCCAA
TCTGATGAGCAATATGCTTGCCATCTTGTTGCCAGAATCAGCTGCCATTGCTGAACTGAGACA
AGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCAT
TCTGGAGTGACATGATGGACTCCGC

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAAGCT
GAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGACCCA
GGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCGGGGACCGCTTCGGCTG
AAGCATTTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCC
TTGCACACCTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAAT
TTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTA
CAGAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCA
TTCGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCT
AACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCCGGAGGTGGG
GCGCCGCTGGGGCCGGCCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCCGCGACCGAGCGTGCGG
ACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGGCCGCGGCTGGGGATTCTT
GTTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGCAGCCCCGGAGA
CAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGTACCTGTGATGTTGAA
ACCATTTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAACTTCTTGAAAGTGACTA
CTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCCTTCTGGAATGACATCAGCCAGTGTG
GAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAAGTTCCTGATGGAATTAAATCTGCG
AGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGAAGAATGTGAACAAGCTGAACGACTTGG
AGCAGTGGATGAATCTCTGAGTGAGGAAACACAGAAGGCTGTTCTTCAGTGGACCAAGCATGATG
ATTCTTCAGATAACTTCTGTGAAGCTGATGACATTCAGTCCCCTGAAGCTGAATATGTAGATTTG
CTTCTTAATCCTGAGCGCTACACTGGTTACAAGGGACCAGATGCTTGGAAAATATGGAATGTCAT
CTACGAAGAAAACCTGTTTTAAGCCACAGACAATTAAGACCTTTAAATCCTTTGGCTTCTGGTC
AAGGGACAAGTGAAGAGAACACTTTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAGAGCA
TTCTACAGACTTATATCTGGCCTACATGCAAGCATTAAATGTGCATTTGAGTGCAAGATATCTTTT
ACAAGAGACCTGGTTAGAAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATG
GAATTTTGACTGAAGGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTTCTCTACTTAATAGAA
CTAAGGGCTTTATCCAAAGTGTTACCATTCTTCGAGCGCCAGATTTTCAACTCTTTACTGGAAA
TAAAATTCAGGATGAGGAAAACAAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTT
CTTTGCATTTTGTATGAGAATTCATTTTTTGTCTGGGGATAAAAAAGAAGCACACAACTAAAGGAG
GACTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTCTG
TCTGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAGAAAT
TGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACCAGACAAGAAATAGTA
TCATTATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAACCTTCAGGAACCT
GTTACAGAATATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTTCTGGACAATGGAGGCGAAA
GAGTGGAATTTCAATCAAAGGCATAATAGCAATGACAGTCTTAAGCCAAACATTTTATATAAAGT
TGCTTTTGTAAGGAGAATTATATTGTTTTAAGTAAACACATTTTAAAAATTGTGTTAAGTCTA
TGTATAATACTACTGTGAGTAAAAGTAATACTTTAATAATGTGGTACAAATTTTAAAGTTTAATA
TTGAATAAAAGGAGGATTATCAAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974

<subunit 1 of 1, 468 aa, 1 stop

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAQAQRFCQVSGYLDDCTCDVETIDRFNNYRLFPRL
QKLESDFRYKVNLRPCPFWNDISQCGRDCAVKPCQSDEVDPDGIKSASYKYSEEANNLIEE
CEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLNPERYTGYKGP
AWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFYRLISGLHASINV
HLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLIELRALSQVLPFFERP
DFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAHKLKEDFRLHFRNISRIMD
CVGCFKCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFHLTRQEIVSLFNAFGRISTSVK
ELENFRNLLQNIH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAACTGTTTTTAAGCCACAGACAATTAAAAGACCTTTAA
ATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTTACAGTTGGCTAGAAGGTCTC
TGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAATGTGCATTT
GAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACAACATTACAGAAT
TTNAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGGCTTAAGAACTTGTAT
TTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATTCTTNGAGCGCCAGATTT
TCAACTNTTTTACTGGAAATAAAATTCAGGATGAGGNAAACAAAATGTTACTTTTGGAAATACTTC
ATGAAATCAAGTCATTTTCCTTTGCATTTTGATGAGAATTCATTTTTTTTGCTG

FIGURE 138

CGGACGCGTGCGGCGGACGCGTGCGGCGGACGCGTGCGGTGGGAGGGGGCAGGATGGGAGGGAAAGT
GAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAACCGATCA
GGCCATGGAACTCCCCCTTCGTCACTCACCTGTTCTTGCCCCCTGGTGTTCTTGACAGGTCTCTGCTC
CCCCTTTAACCTGGATGAACATCACCCACGCCTATTCACAGGGCCACCAGAAGCTGAATTTGGAT
ACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGG
CCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGGGGCCCACAATGCCCCATG
TGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTCATCCTGCTGTGAATATGCACC
TGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTTCATGGTGAGCTTAAGGAGAGGGTGGT
GGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAGAGAAGTGTGGTAAGGGAAAATGGTCTGTGTG
GAGGGGTCAAGGAGTTAAAAACCCTAGAAAGCAAAAGGTAGGTAATGTCAGGGAGTAGTCTTCAT
GCCTCCTTCAACTGGGAGCATGTTCTGAGGGTGCCCTCCCAAGCCTGGGAGTAATAATTTCCCCC
ATCCCCAGGCCTGTGCCCCCTCTCTGGTCTCGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATA
TGTGCCCCGTGTGGATGCTTCATTCCAGCCTCAGGGAAGCCTGGCACCCACTGCCCAACGTGAGCC
AGAGGAAGGCTGAGTACTTGGTTCCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAGCG
GTATGATGCCTGGCAAAGGGCCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAAGCT
CCATGTTTTCCTAACAGATTCAGACTCCTGGCCAGGTGTGGTGGCCCACACCTGTAATTCTAGCAC
TTTGGGAGGCCAAGGTGGGCAGATCACTTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG
TGAAACTCCATCTCTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
GTAATCTCATCTACTCGGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTG
CGGTGAGCCAAGATTGTGCCTCTGCACTCTAGCGTGCGGTGACAGAGTAAGCGAGACTCCATCTCA
AAAATAATAATAATAATAATTTCAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAAC
TCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAGGTTT
GAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPWDGP
SGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCTTCCAACTTTATTTTTAGCTTGCCATTGGGAGGG
GGCAGGATGGGAGGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCA
TACTGGACAGAAACCGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTTGCCCCTGGTG
TTCCTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCTATTCCCAGGGCC
ACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGG
TGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGG
GGGGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTCA
TCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGG

FIGURE 141

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTGGGC
AGAAAGGAGGGTGCTTCGGAGCCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACAATTGAG
GCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCTGAGATGGAC
AGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCAAATGCAGACTTT
CACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCTACGCATTGATTCCAT
GTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAAC
ATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGAAACAGTGACTATTCTGTGCGA
ATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCTGGATCCCCCAGCAGCTGGTGCTCAC
TCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTCGT
GTCAGGGCCACATTGGGCTCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAA
CTCAACCATCCCTTACCCGACCTGGGATGGAGATCACCAAAGATGGCTTCCACCTGGTTATTGAGC
TGGAGGACCTGGGGCCCCAGTTTGAGTTCCTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAG
GAACATGTCAAATGGTGAGGAGTGGGGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGC
TGCATACTGTGTGAAGGCCCAGACATTTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGA
CAGAATGTGTGGAGGTGCAAGGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGC
TTCATGCTGATCCTTGTGGTCGTGCCACTGTTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTC
CTGTTGCCCCGTGGTGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCAGAGTTAATCA
GCTGCAGAAGGGAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACCTCCTCAGG
GCCTGGATCTCATAGGTTTTCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGG
AAACCATGAGGGGACAAGTTGTGTTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAA
GAGCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTTGTCTAACAGAACCT
GACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACCCTGG
GAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACCTACACAC
CTGCTAAACACACACACACAGAGTCTCTCTATATATACACACGTACACATAAATACACCCAGC
ACTTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATTGAGTGTCTGAGAG
CAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGTGGCTTGGAGAGCCAC
TTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGGTGTGAGTTCACTTCAAGC
CCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGTAGGTGACCTGGAGGAAGGTCAC
AGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATCCATGAACTACTGTAAAGTGTGACA
GTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGTGTGCAATGCGACGAGAATGCAGAAGTCA
GTAACATGTGCATGTTTGTGTGCTCCTTTTTTCTGTTGGTAAAGTACAGAATTCAGCAAATAAA
AAGGGCCACCCTGGCCAAAAGCGGTAAAAA

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSLSFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLMWSPIAPGETVY
YSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAWSILKHP
FNRNSTILTRPGMEITKDGPHLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSGGIPVHLETM
EPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPVLALFAFVGFMILVVVPLFVWKMGRLL
LQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTTCCT
GGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCATTCTGA
AGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGA
GTCTACCAAATGCAGACTTTTACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTT
TTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACC
TCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGAA
ACAGTGTACTATTCTGTGGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCTGGAT
CCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATCACGGCCA
CTGTGCCATACAACCTTTGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTGGAGCATCCTG
AAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGATCACCAAAGATGG
CTTNCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGA
GGAGGGGCGAACCCCTTGCGGCGCAAGGGGTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTC
GAGAAAAGAGAGGCCCAATATGACCCACATACTCAATATGGACGAANTGCTATTGTCCACCTGTT
TGAGTGGCGCTGGGTTGAT

FIGURE 144

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGAGGC
AGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGAGGAGGT
GAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAGGAGATG
CGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAACGGAGAGGAGGTG
TGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGAGTAGGAAGATCAGGA
GCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGGAAAGACACAGAGAGACGG
GAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTGGCTGCTTTGGCATTGTTGGGAA
CTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCTGGAGGGACAGGGTCCCAGAAGGAG
GGGACAGAGGAGCTGAGAGAGGGGGGCAGGGCGTTGGGCAGGGGTCCCTCGGAGGCCTCCTGGGG
ATGGGGGCTGCAGCTCGTCTGAGCGCCCCCTCGAGCGCTGGTACTCTGGGCTGCACTGGGGGCAGC
AGCTCACATCGGACCAGCACCTGACCCCGAGGACTGGTGGAGCTACAAGGATAATCTCCAGGGAA
ACTTCGTGCCAGGGCCTCCTTTCTGGGGCCTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGG
AAGCGGCAGAGCCCCGTGGATGTGGAGCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCCATT
AAGGCTCAGCACTGGAGGAGAGAAGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCT
TCCTGCCTGCACCCCGACCTGTGGTCAATGTGTCTGGAGGTCCCTCCTTTACAGCCACCGACTC
AGTGAACTGCGGCTGCTGTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACCA
GGGCTTCTCTGCTGAGGTGCAGCTCATTCACTTCAACCAGGAACCTCTACGGGAATTTACAGCGCTG
CCTCCCGCGGCCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAAC
CCATTCTCAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTA
CTTTCTTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGG
GCTCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAAT
ATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCTTCCA
GAGCCTCAGCGGTAACAGCCGGCCCCCTGCAGCCCTTGGCCACAGGGCACTGAGGGGCAACAGGG
ACCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCAACTACCGCCTGCATGTGGATGGTGTCCCC
CATGGTCGCTAGACTCCCCTTCGAGGATTGCACCCGCCCGTCCTAAGCCTCCCCACAAGGCGAG
GGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLCAVG
KRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPLYSHRL
SELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSLFVNVASTSN
PFLSRLLNRTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSETVTWILIDRALN
ITSLQMHSLRLLSQNPSPQIFQSLSGNSRPLQPLAHRALRGNRDPRHPERRCRGPNYRLHVDGVP
HGR

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCCGCCAGCCTCCGCCGCCGAGCCTCGTT
CGTGTCCTCCCGCCCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCACCCTGGCAGACTAACGAAGCAGCT
CCCTTCCCACCCCAACTGCAGGTCTAATTTTGGACGCTTTGCCTGCCATTTCTTCCAGGTTGAGGGAGCCGCAGAGGC
GGAGGCTCGCGTATTCTCGCAGTCAGCACCCACGTCGCCCCCGACGCTCGGTGCTCAGGCCCTTCGCGAGCGGGGCT
CTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGCGCGTCCGGTTTGGCTCACCTCTCCCAGGAA
ACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGGAGATTTTCTGGGAAATCCTGAGGTCAATCATT
ATGAAGTGTACCGCGCGGAGTGCGCTCAGAGTAACCACAGTGCTGTTTCATGGCTAGAGCAATTCAGCCATGGTGGTT
CCCAATGCCACTTTATTGGAGAACTTTTGGAAAAATACATGGATGAGGATGGTGAGTGGTGGATAGCCAAACAACGA
GGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCATAATAAATTACGAAGTCAGGTGTATCCA
ACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGATCTGCAGAATCCTGGGCTGAAAGTTGCTTG
TGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAATTTGGGAGCACACTGGGGAAGATATAGGCCCCCG
ACGTTTTCATGTACAAATCGTGATGATGAAGTGAAAGACTTTAGCTACCCATATGAACATGAATGCAACCCATATTGT
CCATTCAAGGTGTTCTGGCCCTGTATGTACACATTATACACAGGTCGTGTGGGCAACTAGTAACAGAATCGGTTGTGCC
ATTAATTTGTGTCTAATCATGAACATCTGGGGGCGAGATATGGCCCAAAGCTGTCTACCTGGTGTGAATTACTCCCCA
AAGGGAACTGGTGGGCCATGCCCCCTTACAAACATGGGCGGCCCTGTCTGCTTGCCACCTAGTTTTGGAGGGGGC
TGTAGAGAAAATCTGTGCTACAAAGAAGGGTCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATGAAATAGAA
CGACAGCAGTCACAAAGTCCATGACACCCATGTCCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCA
CAGCAAATGTCCCAAATGTTTCTTGTGAAGTAAGATTAAGAGATCAGTGCAAAGGAACAACCTGCAATAGGTACGAA
TGTCTGCTGGCTGTTTGGATAGTAAAGCTAAAGTTATTGGCAGTGATACATTATGAAATGCAATCCAGCATCTGTAGA
GCTGCAATTCATTATGGTATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTT
ATCAAGTCCAATAGAAATGGTATTCAACAATTTGGCAAATATCAGTCTGCTAATTCCTTCACAGTCTCTAAAGTAACA
GTTCAAGGCTGTGACTTGTGAAACAACCTGTGGAACAGCTCTGTCCATTTTATAAGCCTGCTTCACATTGCCCAAGAGTA
TACTGTCTCTGTAACGTATGCAAGCAAATCCACATTATGCTCGTGTAATTGGAACCTCGAGTTTATTCTGATCTGTCC
AGTATCTGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCACGGTGGTTATGTTGATGTAATGCCTGTGGACAAA
AGAAAGACCTACATTGCTTCTTTTTCAGAATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATT
AGAGTGTGTTGCTGTTGTGTGAAACTGAATACTTGGAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGAATT
TTGTATAAACTGTAACTTACTGTACAGAGTACATCAACTATTTTTCAGCCCCAAAAGGTGCCAAATGCATATAAATC
TTGATAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAATGAAAATATAATGGTTTTAGAAA
TCCTGTGTTAAATATTGCTATATTTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATTGTTCTACGTTT
ATATATTATATGGTGCTTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATGGCCCTCAGAAAATCA
TCTAGTGCAATTTAAAAATAATCGACTCTAAACTGAAAGAAACCTTATCACATTTTCCCGAGTTCAATGCTATGCCAT
TACCAACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTTCTGTTAATTTAGGCATATAGAA
TATTAAATCTGATATTGCACTTCTTATTTTATATAAAATAATCCTTTAATATCCAAATGAATCTGTTAAATGTTTG
ATTCCTTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGTATGAAAACATTCTAGTGATCATGTAG
TAAATGTAGGGTTAAGCATGGACAGCCAGAGCTTTCTATGTACTGTAAAAATTGAGGTCACATATTTTCTTTTGTATC
CTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTTGAACAAAGATGAACTAATGTATTACATTACCAT
TGCCACTGATTTTTTTTAAATGGTAAATGACCTTGTATATAAATATTGCCATATCATGGTACCTATAATGGTGATATA
TTTGTCTTATGAAAAATGATTGTGCTTTGATACTAAAAATCTGTAAAAATGTTAGTTTTGGTAATTTTTTTCTGCT
GGTGGATTACATATTAAATTTTTTCTGCTGGTGGATAAACATTAATAATCATGTTTCAAAAAAAAAAAAAA

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAI PAMVVPNATLLEKLLEKYMDEEDGEWWIAKQRGKRAITDNDMQSI
LDLHNKLRSQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHWGRYRPP
TFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQVVWATSNRIGCAINLCHNMNIWGQI
WPKAVYLVLCNYSKGNWWGHAPYKHGRPCSACPPSFGGGCRENL CYKEGSDRYPPREEETNEIE
RQQSQVHDTHVRTRSDSSRNEVIS AQQMSQIVSCEVRLRDQCKGTTCNRYECPAGCLDSKAKVI
GSVHYEMQSSICRAAIHYGII DNDGGWVDITRQGRKH YFIKSNRNGIQTIGKYQSANSFTVSKVT
VQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIGTRVYSDLSSICRAAVHAGVVR
NHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKA FRVFAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCGCAG
CCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGTGGCGTC
TCCGGGCCCGCGCTCCGACGGGGCCAGCGCCCTCCCCATGTCCCTGCTCCCACGCCGCGCCCCCTCC
GGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGTACACCGCGCGTG
TGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTACAGCGACGTGAAGAAG
CTGGAAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTATCATCACCACCAAGAGCGT
GTCCAGGTACCGAGGTCAGGAGCACTGCCTGCACCCCAAGCTGCAGAGCACCAAGCGCTTCATCA
AGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAAGAATAGGGTGAAAAACCTCAGAAG
GGAAAACCTCAAACCAGTTGGGAGACTTGTGCAAAGGACTTTGCAGATTAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCCTTTCTTTCTCACAGGCATAAGACACAAATTATAT
ATTGTTATGAAGCACTTTTTTACCAACGGTCAGTTTTTTACATTTTATAGCTGCGTGCGAAAGGCTT
CCAGATGGGAGACCCATCTCTCTTGTGCTCCAGACTTCATCACAGGCTGCTTTTTTATCAAAAAGG
GGAAAACCTCATGCCTTTCTTTTTTAAAAAATGCTTTTTTGTATTTGTCCATACGTCACTATACAT
CTGAGCTTTATAAGCGCCCGGGAGGAACAATGAGCTTGGTGGACACATTTTATTGCAGTGTTGCT
CCATTCCCTAGCTTGGGAAGCTTCCGCTTAGAGGTCTTGGCGCCTCGGCACAGCTGCCACGGGCTC
TCCTGGGCTTATGGCCGGTCACAGCCTCAGTGTGACTCCACAGTGGCCCCCTGTAGCCGGGCAAGC
AGGAGCAGGTCTCTCTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTT
CATTCCTCCCTGGTTAATTTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGAC
AAATGATCCTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTTACAGGTT
CAATATTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAA
CTTAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAAA
ACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATATTTGT
GATTCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCTCTTTTCGAAAGCTAAGATGACCATG
CGCCCTTTCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCCAGTATATGC
CGCATTGTACTGCTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCATTGCATGCAGGTTT
CATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAAAAAAAAAAAAAAA

FIGURE 149

MSLLPRRAPPVSMRLLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPHCEE
KMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNWNEKRRVYEE

Signal sequence:

amino acids 1-34

FIGURE 150

CCCCAGGGACTGCTATGGCTTCCTTTGTTGTTTACCCCGGTCTGCGTCATGTTAAACTCCAATGTCTCTCTGTGGTT
AACTGCTCTTGCCATCAAGTTCACCTCATTTGACAGCCAAGCACAGTATCCAGTTGTCAACACAAATTATGGCAAAAT
CCGGGGCCTAAGAACACCGTTTACCCAATGAGATCTTGGGTCCAGTGGAGCAGTACTTAGGGGTCCCCATGCCTCACC
CCCCACTGGAGAGAGGCGGTTTCAGCCCCGAGAACCCCGTCTCTGGACTGGCATCCGAAATACTACTCAGTTTGC
TGCTGTGTGCCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCATCTGGTTTACCGCCAATTTGGA
TACTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCTTTACTTAAACATCTACGTGCCCACGGAAGATGGAGC
CAACACAAAGAAAAACGCAGATGATATAACGAGTAATGACCGTGGTGAAGACGAAGATATTCATGATCAGAACAGTAA
GAAGCCCGTCATGGTCTATATCCATGGGGGATCTTACATGGAGGGCACCGGCAACATGATTGACGGCAGCATTTTGGC
AAGCTACGGAAACGTCATCGTGATCACCATTAACTACCGTCTGGGAATACTAGGGTTTTTAAGTACCGGTGACCAGGC
AGCAAAAGGCAACTATGGGCTCCTGGATCAGATTCAAGCACTGCGGTGGATTGAGGAGAATGTGGGAGCCTTTGGCGG
GGACCCCAAGAGAGTGACCATCTTTGGCTCGGGGGCTGGGGCCTCCTGTGTGTCAGCCTGTTGACCCTGTCCCCTACTC
AGAAGGTCTCTTCCAGAAGGCCATCATTCAGAGCGGCACCGCCCTGTCCAGCTGGGCAGTGAACCTACCAGCCGGCCAA
GTACACTCGGATATTGGCAGACAAGGTGCGCTGCAACATGCTGGACACCACGGACATGGTAGAATGCCTGCGGAACAA
GAAGTACAAGGAGCTCATCCAGCAGACCATCACCCCGGCCACCTACCACATAGCCTTCGGGCCGGTGATCGACGGCGA
CGTCATCCCAGACGACCCCGAGATCCTGATGGAGCAAGGCGAGTTCTCTCACTACGACATCATGCTGGGCGTCAACCA
AGGGGAAGGCTGAAGTTCTGTGGACGGCATCGTGGATAACGAGGACGGTGTGACGCCCAACGACTTTGACTTCTCCGT
GTCCAACCTTCGTGGACAACCTTTACGGCTACCCCTGAAGGGAAGACACTTTGCGGGAGACTATCAAGTTTCATGTACAC
AGACTGGGCGGATAAGGAAAACCCGAGACGCGCGGAAAACCCCTGGTGGCTCTCTTTACTGACCACAGTGGGTGGC
CCCCGCGGTGGCCGCGGACCTGCACGCGCAGTACGGCTCCCCACCTACTTCTATGCCTTCTATCATCACTGCCAAAG
CGAAATGAAGCCAGCTGGGCAGATTTCGGCCCATGGTGATGAGGTCCCTATGTCTTCGGCATCCCCATGATCGGTCC
CACCGAGCTCTTCAGTTGTAACCTTTTCCAAGAACGACGTCATGCTCAGCGCCGTGGTCATGACCTTGGACGAACCT
CGCCAAAACCTGGTGATCCAAATCAACCAGTTCCTCAGGATACCAAGTTCATTACACAAAACCCAAACCGCTTTGAAGA
AGTGGCCTGGTCCAAGTATAATCCCAAAGACCAGCTCTATCTGCATATTGGCTTGAAACCCAGAGTGAGAGATCACTA
CCGGGCAACGAAAGTGGCTTTCTGGTTGGAACCTCGTTCCTCATTTGCACAACCTTGAACGAGATATTCCAGTATGTTTC
AACAACCACAAAGTTCTCCACCAGACATGACATCATTTCCCTATGGCACCCGGCGATCTCCCGCCAAGATATGGCC
AACCACCAAACGCCCAGCAATCACTCCTGCCAACAATCCCAAACACTCTAAGGACCCCTCACAAAACAGGGCCTGAGGA
CACAACCTGTCTCTATTGAAACCAAACGAGATTATTCCACCGAATTAAGTGTACCATTTGCCGTGGGGCGTCTGCTCCT
CTTCTCTCAACATCTTAGCTTTTGCGGCGCTGTACTACAAAAGGACAAGAGGCGCCATGAGACTCACAGGCGCCCCAG
TCCCCAGAGAAACACCACAAATGATATCGCTCACATCCAGAACGAAGAGATCATGTCTCTGCAGATGAAGCAGCTGGA
ACACGATCACGAGTGTGAGTGTGCTGCAGGCACACGACACACTGAGGCTCACCTGCCCGCCAGACTACACCTCACGCT
GCGCCGGTGGCCAGATGACATCCCACTTATGACGCCAAACACCATCACCATGATTCCAAACACACTGACGGGGATGCA
GCCTTTGCACACTTTTAACACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCACGGACATTCCACCACTAGAGT
ATAGCTTTGCCCTATTTCCCTTCCTATCCCTCTGCCCTACCCGCTCAGCAACATAGAAGAGGGAAGGAAAGAGAGAAG
GAAAGAGAGAGAGAAAGAAAGTCTCCAGACCAGGAATGTTTTGTCCCACTGACTTAAGACAAAATGCAAAAAGGCA
GTCATCCCATCCCGGCAGACCTTATCGTTGGTGTGTTTTCCAGTATTACAAGATCAACTTCTGACCTGTGAAATGTGA
GAAGTACACATTTCTGTAAAAATAACTGCTTTAAGATCTCTACCACTCCAATCAATGTTTAGTGTGATAGGACATCAC
CATTTCAAGGCCCGGGTGTGTTTCCAACGTCATGGAAGCAGCTGACACTTCTGAAACTCAGCCAAGGACACTTGATATT
TTTTAATTACAATGGAAGTTTAAACATTTCTTTCTGTGCCACACAATGGATGGCTCTCCTTAAGTGAAGAAAGAGTCA
ATGAGATTTTGCCAGCACATGGAGCTGTAATCCAGAGAGAAGGAAACGTAGAAATTTATTATTAAGAATGGACTG
TGCAGCGAAATCTGTACGTTCTGTGCAAGAGGTGTTTTGCCAGCCTGAACATATTTAAGAGACTTTGT

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASPPTG
ERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQNEDECLYL
NIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMIDGSILASYG
NVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKRVTIFGSGAGASC
VSLTTLSHYSEGLFQKAI IQSGTALSSWAVNYQPAKYTRILADKVGCNMLDTTDMVECLRKNKYK
ELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLGVNQGEGLKFVDGIVDNEDG
VTPNDFDFSVS NFVDNLYGYPEGKDTLRETIKFMYTDWADKENPETRRKTLVALFTDHQWVAPAV
AADLHAQYGSPTYFYAFYHHCQSEMKPSWADSAHGDEVYPYVFGIPMIGPTELFSCNFSKNDVMLS
AVVMTYWTNFAKTGDPNQVPVQDTKFIHTKPNRFEEVAWSKYNPKDQLYLHIGLKPRVRDHYRAT
KVAFWLELVPHLHNLNEIFQYVSTTTKVPPPDMSFPYGTRRSPAKIWPTTKRPAITPANNPKHS
KDPHKTGPEDTTVLIETKRDYSTE SVTI AVGASLLFLN ILAFAALYYKKDKRRHETHRRPSPQR
NTTNDIAHIQNEEIMSLQMKQLEHDHECESLQAHDTLRLTCPPDYTLTLRRSPDDIPLMTPNTIT
MIPNTLTGMQPLHTFNTFSGGQNSTNLPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAAGATGGCGGCGACTCTGGGACCCCTTGGGTCTGTGGCAGCAGTGGCGGCGATGTTTGTCTGG
CTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTGGGGTCTGGGCAGGGGCCACAGCAA
GTCGGGGCGGGTCAAACGTTTCGAGTACTTGAAACGGGAGCACTCGCTGTGCAAGCCCTACCAGGG
TGTGGGCACAGGCAGTTCCCTCACTGTGGAATCTGATGGGCAATGCCATGGTGATGACCCAGTATA
TCCGCCTTACCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGGGTGCCATGTTTCCTG
AGAGACTGGGAGTTGCAGGTGCACTTCAAAATCCATGGACAAGGAAAGAAGAATCTGCATGGGGA
TGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGCCTGTGTTTGGAAACATGGACA
AATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAATGAGGAGAAGCAGCAAGAGCGGGTA
TTCCCCTACATCTCAGCCATGGTGAACAACGGCTCCCTCAGCTATGATCATGAGCGGGATGGGCG
GCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCCGCAATCTTCATTACGACACCTTCCTGGTGA
TTCGCTACGTCAAGAGGCATTTGACGATAATGATGGATATTGATGGCAAGCATGAGTGGAGGGAC
TGCATTGAAGTGCCCGGAGTCCGCCTGCCCCGCGGCTACTACTTCGGCACCTCCTCCATCACTGG
GGATCTCTCAGATAATCATGATGTCAATTCCTTGAAGTTGTTTGAAGTACAGTGGAGAGAACCC
CAGAAGAGGAAAAGCTCCATCGAGATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAG
ATGACAGCTCCACTGCCGCCCCCTGAGTGGCCTGGCCCTCTTCCTCATCGTCTTTTTCTCCCTGGT
GTTTTCTGTATTTGCCATAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAA
AGCGCTTCTACTGAGCCCTCCTGCTGCCACCACTTTTGTGACTGTCACCCATGAGGTATGGAAGG
AGCAGGCACTGGCCTGAGCATGCAGCCTGGAGAGTGTCTTGTCTCTAGCAGCTGGTTGGGGACT
ATATTCTGTCACTGGAGTTTTGAATGCAGGGACCCCGCATTCCTCATGGTTGTGCATGGGGACATC
TAACTCTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCTTTCCCTGCAGT
CCTTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAAATCACAGAACA
GAATTTCATAGCCCAGGCTGCCGTGTTGTTTGAAGGAGGCTTCTACTTCAGTTTTGAATC
CACAAAGAATTAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTTCGTTGGGTTTTGCATT
TGACCCAACCCCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAAACTTCTTCCCTGC
CTTACCTTCCCTTTCACTCCATTCAATTGTCCTCTCTGTGTGCAACCTGAGCTGGGAAAGGCATTTG
GATGCCTCTCTGTTGGGGCCTGGGGCTGCAGAACACACCTGCGTTTCACTGGCCTTCATTAGGTG
GCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCATGGGTCTTGGGTCTATTGGC
ATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTGAAGTTTGGCTAAAGGTTGGTGT
AAAAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATGGATTAGCTGTGCAACTGACCAGCTC
CAGGTTTGTATCAAAACCAAAAGCAACATTTGTTCATGTGGTCTGACCATGTGGAGATGTTTCTGGAC
TTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGTTACGATTTTTTGAATCCCACTTTGAGTGTCTG
AAAGTGTAAGGAAGCTTTCTTCTTACACCTTGGGCTTGGATATTGCCAGAGAAAGAAATTTGGCT
TTTTTTTTCTTAATGGACAAGAGACAGTTGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACC
CTCATCATCTGTGCCTGGAAGAGTTCACTGTCAATTGAGCAGCACAGCCTGAGTGTGCTGGCCTCTGT
CAACCCTTATTCCACTGCCTTATTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGAT
TAAATCAGTTACAGGCCAGAGTCTCCTTGGAGGGCCTGGAACCTCTGAGTCCTCCTATGAACCTCT
GTAGCCTAAATGAAATCTTAAATACCCGATGGAACCAAAAAAAAAAAAAAAAAAGGGCGGCCCG
GACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPPQVGAGQTFEYLLKREHSLSKPYQGVG
TGSSSLWNLGMNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKEHGQKKNLHGDGL
AIWYTKDRMQPGPVFGNMDKFFVGLGVFVDTPNEEKQQERVPYISAMVNNGSLSYDHERDGRPT
ELGGCTAIVRNLYHDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRGYFSGTSSITGDL
SDNHDVISLKLFEFTVERTPEEEKLHRDVFLPSVDNMKLPENTAPLPPLSGLALFLIVFFSLVFS
VFAIVIGIILYNKWQEQRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCTCTG
ATGGTGGGTCGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAAATCAGTTGTCTTCATTCAAGAATTGGGGT
CTGGCTCAGAAATTCCTGCAGCTGGTGAAAATCTGTTTTCTAGAAGAGGTTAATTAATGCCTGCAGTCTGACATGTTT
CCGATTTGAGGTGAAACCATGAAGAGAAAAATAGAATACTTAATAATGCTTTTCCGCAACCGCTTCTTGCTGCTGCTGG
CCCTGGCTGCGCTGCTGGCCTTTGTGAGCCTCAGCCTGCAGTTCCTCCACCTGATCCCGGTGTGACTCCTAAGAATG
GAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCGACCCCTGTGACGGAGCCCCCTGTGACAGACCCCGTTTATGAAG
CTCTTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTCATGCCCCGCATCTTTAAGCTGGTCT
CAGTGCATGTGTTTCATTTCGCCACGGAGACAGGTACCCACTGTATGTTCATTCCCAAAACAAAGCGACCAGAAATTGACT
GCACTCTGGTGGCTAACAGGAAACCGTATCACCCAAAACCTGGAAGCTTTTCATTAGTCACATGTCAAAAGGATCCGGAG
CCTCTTTTCGAAAGCCCCCTTGAACCTCTTGCCTCTTTACCCAAATCACCATTGTGTGAGATGGGAGAGCTCACACAGA
CAGGAGTTGTGCAGCATTTGCAGAACGGTCAGCTGCTGAGGGATATCTATCTAAAGAAACACAACTCCTGCCCAATG
ATTGGTCTGCAGACCAGCTCTATTTAGAGACCACTGGGAAAAGCCGGACCTACAAAGTGGGCTGGCCTTGCTTTATG
GCTTTCTCCAGATTTTGTACTGGAAGAAGATTTATTTTCAGGCACCAGCCAAGTGCCTGTTCTGCTCTGGAAGCTGCT
ATTGCCCGGTAAGAAACAGTATCTGGAAGAGGAGCAGCGTCGTACCTCCTACGTTTGAAAAACAGCCAGCTGG
AGAAGACCTACGGGGAGATGGCCAAGATCGTGGATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCA
TGCTCTGCCACTTCTGCCACAATGTGAGCTTTCCCTGTACAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAA
TTAAGACCCATCAGATCGAGGATGAAAGGGAAAGACGGGAGAAGAAATGTACTTCGGGTATTCTCTCCTGGGTGCC
ACCCCATCCTGAACCAAAACCATCGGCCGGATGCAGCGTGCCACCGAGGGCAGGAAAGAAGAGCTCTTTGCCCTCTACT
CTGCTCATGATGTCACTCTGTCAACAGTTCTCAGTGCCTTGGGCCCTTTCAGAAGCCAGGTTCCCAAGGTTTGCAGCCA
GGTTGATCTTTGAGCTTTGGCAAGACAGAGAAAAGCCAGTGAACATTCCTGCCGATTCTTTACAATGGCGTCGATG
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGCACCT
GCTTTGTGAAAAGGGCATGTTTGTAGCCCTGGGTGGCAGTGGTACAAATATTATGATGCATGTACAGGGAAGGAT
TCTAAAAGGTATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTTGTC
TGTTACTAAGGTTAGAAGATTATTGCTTTTTTAAAGGCTAAATATTGTTTGTGGGAACCAAGATGGTTGGGGTTGAAC
AGTAAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGGCCAGTTCACAGAGGAATAGAAGTACTT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGCACCT
TCTGGCCTGCCCATGTTACTATGTGATGGAACCAGCACACCTCAACCAAAATTTTTTAAATCTTAGACATTTTTACC
TTGTCTTGTAAAGATTTCTTGAAGTGATTTATCTAAAATAAAGGTTGGCAAATTTTTCTGTAAAGGGCCAGATTG
TAAATATTTAGACTGTGTGGACCAAAAGGCCACATACAGTCTCTGTACATAACTACTCAACTCTGTTTCTGAAGCAGG
AAAGCCACCACAGACAGTACATAAAGGAATATGTGTAGCTGGGTTCCAGGCCAGACAAAACAGATGGTGACCAGACT
TGGCCCCCTGGGCTGTAGTTTGTGCTGACCCCTCATCTAAAAAATAGGCTATACTACAATTGCACTTCCAGCACTTTGAGA
ACGAGTTGAATACCAAGAATTATTCAATGGTTCTCCAGTAACTTCTGCTAGAAACACAGAATTTGGTCTGTATCTGA
CACTAGAACAAAATTTGAGGGTAAATAAATGAAATGAAATGAAATCATAGAAAATGATTAGAAGAATACTTGATG
TTTATGATGATTGTGGTACAAGATAGTTTTAAGTATGTTCTAAATATTTGTCTGCTGTAGTCTATTTGCTGTATATGC
TGAAATTTTTGTATGCCATTTAGTATTTTTATAGTTTTAGGAAAAATATTTCTAAGACCAGTTTTAGATGACTCTTATT
CCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAGGAGGCTAGAAGATGAATTCAGGCACCTTTCTTCC
AATAAACTAATTATGGCTCATTCCCTTTGACAAGCTGTAGAACTGGATTCATTTTTAAACATTTTCATCAGTTTCA
AATGGTAAATTTCTGATTGATTTTTAAATGCGTTTTTGGAAAGAACTTTGCTATTAGGTAGTTTACAGATCTTTATAAGG
TGTTTTATATATTAGAAGCAATTATAATTACATCTGTGATTTCTGAACTAATGGTGCTAATTCAGAGAAATGGAAAGT
GAAAGTGAGATTCTCTGTTGTGTCATCGGCATTCCAACTTTTCTCTTTTGTCTTGTCCAGTGTTGCATTTGAATATGTC
TGTTTCTATAAAATAAATTTTTTAAAGATAA

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329

><subunit 1 of 1, 480 aa, 1 stop

><MW: 55240, pI: 9.30, NX(S/T): 2

MLFRNRFLLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDPVYEAL
LYCNIPSVASERSMEGHAPHHFKLVSVHVFIRHGDYPLYVIPKTKRPEIDCTLVANRKPYPHPKLE
AFISHMSKSGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLKKHKLLPNDW
SADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCPVRNQYLEKEQRR
QYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSM LCHFCHNVSFPCTRNGCVDMEHFKVIK
THQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELFALYSAHDVTLSPVLSALG
LSEARFPRFAARLIFELWQDREKPSEHSVRILYNGVDVTFHTSFCQDHHKRS PKPMCPLNLVRF
VKRDMFVALGGSGTNYDACHREGF

Signal sequence:

amino acids 1-18

FIGURE 156

[illegible]

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPEELEKGSRVGDISRDLGLEPRELAERGVR
IIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVRDINDNA
PYFRESELEIKISENAATEMRFPPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGADGSKYPELV
LKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYRASVPENLALGTQ
LLVFNATDPDEGVNAEVRYSFRYVDDKAAQVFKLDCNSGTISTIGELDHEESGFYQMEVQAMDNA
GYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVDQDSEENGQVICFIQGNL
PFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLSTETHISLNVADTNDNPPVFPQA
SYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTIQGASLSSYVSINSDTGVLYALSSF
DYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDNAPEILYPALPTDGSTGVELAPRSAEPG
YLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVGLHTGEVRTARALLDRDALKQSLVVAVQDHG
QPPLSATVTTLTVAVADSIPQVLADLGSLESPANSETSDLTLVAVAAVSCVFLAFVILLALLR
LRRWHKSRLQASGGGLTGAPASHFVGVDGVQAFQTYSHEVSLTTDSRKSHLIFPQPNYADMLV
SQESFEKSEPLLLSGDSVFSKDSHGLIEVSLYQIFFLFFFNCSVSQAGVQRYDHSSLRPQTPRLK
QLSHLCLRCNRDYRCKPPTVCLSIYLSIYLSIYLSIYLLLSCTDGSILTPVIPVLWEAEAGGSPEV
GSLRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAGGCT
CCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAATCAGTA
GGTGACCCCGCCCCCTGGATTCTGGAAGACCTCACCAATGGGACGCCCCCGACCTCGTGCGGCCAAG
ACGTGGATGTTCTTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGCACAGGAGGACAA
GGTGCTGGGGGGTTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGGCCTTGTTCCAGGGCC
AGCAACTACTCTGTGGCGGTGTCCTTGTTAGGTGGCAACTGGGTCTTACAGCTGCCCCTGTAAA
AAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAATAAAGATGGCCCAGAGCAAGA
AATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACAGCAGCGATGTGGAGGACCACAACC
ATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGGGTCCAAAGTGAAGCCCATCAGC
CTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTCAGGCTGGGGCACTGTCACCAG
TCCCCGAGAGAATTTTCTGACACTCTCAACTGTGCAGAAAGTAAAAATCTTTCCCCAGAAGAAGT
GTGAGGATGCTTACCCGGGGCAGATCACAGATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCT
GACACGTGCCAGGGCGATTCTGGAGGGCCCCCTGGTGTGTGATGGTGCCTCCAGGGCATCACATC
CTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACC
TGGACTGGATCAAGAAGATCATAGGCAGCAAGGGCTTGATTCTAGGATAAGCACTAGATCTCCCTT
AATAAACTCACAACCTCTCTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVLVGG
NWLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLLQLRDQA
SLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFDTLNCAEVKIFPQKKCEDAYPGQITDG
MVCAGSSKGADTCQGDSSGGLVCDGALQGITSWGSDFCGRSDKPGVYTNICRYLDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

GGCGCCGGTGACACGGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCTGCGCGCCCCGGCCCCGCGC
GCCGCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCGCGCCCGCGCCCCG
GCCCAGGTGAGCGCTCCGCCCCGCGCGAGGCCCCGCCCCGGCCCGCCCCCGCCCCGGCCG
GCGGGGGAACCGGGCGGATTCTCGCGCGTCAAACCACTGATCCCATAAACATTATCCTCCC
GGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCGCCCGCCCTCGCCCTGTGCGCCCTGC
GCGCCCTGCGCACCCGCGGCCCGAGCCCAGCCAGAGCCGGGCGGAGCGGAGCGCGCCGAGCCTCG
TCCCGCGGCCGGGCGGGGCCGGGCCGTAGCGGCGGCGCCTGGATGCGGACCCGCGCCGGGGAG
ACGGGCGCCCCGCCCCGAAACGACTTTTCAGTCCCCGAGCGCCCCGCCCCAACCCCTACGATGAAGA
GGCGCTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTGCTGTGGCTGCAGGCCTGGCAGGTGGCA
GCCCCATGCCCAGGTGCGCTGCGTATGCTACAATGAGCCCAAGGTGACGACAAGCTGCCCCCAGCA
GGGCCTGCAGGCTGTGCCCGTGGGCATCCCTGCTGCCAGCCAGCGCATCTTCCTGCACGGCAACC
GCATCTCGCATGTGCCAGCTGCCAGCTTCCGTGCCTGCCGCAACCTCACCATCCTGTGGCTGCAC
TCGAATGTGCTGGCCCCGAATTGATGCGGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGA
CCTCAGCGATAATGCACAGCTCCGGTCTGTGGACCCTGCCACATTCCACGGCCTGGGCGCCTAC
ACACGCTGCACCTGGACCGCTGCGGCCTGCAGGAGCTGGGCCCCGGGGCTGTTCCGCGGCCTGGCT
GCCCTGCAGTACCTCTACCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGA
CCTGGGCAACCTCACACACCTCTTCTGACGGCAACCGCATCTCCAGCGTGCCCCGAGCGCGCCT
TCCGTGGGCTGCACAGCCTCGACCGTCTCCTACTGCACCAGAACCGCGTGGCCCCATGTGCACCCG
CATGCCTTCCGTGACCTTGGCCGCCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCT
GCCCACTGAGGCGCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGG
TGTGTGACTGCCGGGCACGCCCACTCTGGGCCTGGCTGCAGAAGTTCGCGCGCTCCTCCTCCGAG
GTGCCCTGCAGCCTCCCGCAACGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATGACCT
GCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACCGATGAGG
AGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGTACTGGAGCCT
GGAAGACCAGCTTCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTGACAGCCCGCCGGG
CAACGGCTCTGGCCACGGCACATCAATGACTCACCTTTGGGACTCTGCCTGGCTCTGCTGAGC
CCCCGCTCACTGCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTTCCCCACCTCGGGCCCTCGC
CGGAGGCCAGGCTGTTACGCAAGAACCGCACCCGCAGCCACTGCCGTCTGGGCCAGGCAGGCAG
CGGGGGTGGCGGGACTGGTGACTCAGAAGGCTCAGGTGCCCTACCCAGCCTCACCTGCAGCCTCA
CCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCTTGGGCCCTGCTGAACCCCCAGCGGACACA
AGAGCGTGCTCAGCAGCCAGGTGTGTGTACATACGGGGTCTCTCTCCACGCCGCAAGCCAGCCG
GGCGGCCGACCCGTGGGGCAGGCCAGGCCAGGTCCTCCCTGATGGACGCTGCCGCCCGCCACCC
CCATCTCCACCCCATCATGTTTACAGGGTTCCGGCGGCAGCGTTTGTTCAGAACGCCGCTCCCA
CCCAGATCGCGGTATATAGAGATATGCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGG
AATAAAGAGCTCTTTTCTTAAAAAA

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRL LAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRIFLH
GNRISHVPAASFRACRNLTILWLHSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPATFHGLG
RLHTLHLDRCLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLHGNRISVPE
RAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAPLRALQYLRLNDN
PWVCDRCRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCAVATGPYHPITWGRAT
DEEPLGLPKCCQPDAADKASVLEPGRPASAGNALKGRVPPGDSPPGNGSGPRHINDSPFGTLPGS
AEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQAGSGGGGTGDSEGS GALPSLTC
SLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

FIGURE 162

GGAAGTCCACGGGGAGCTTGGATGCCAAAGGGAGGACGGCTGGGTCTCTGGAGAGGACTACTCA
CTGGCATATTTCTGAGGTATCTGTAGAATAACCACAGCCTCAGATACTGGGGACTTTACAGTCCC
ACAGAACCGTCCTCCCAGGAAGCTGAATCCAGCAAGAACAAATGGAGGCCAGCGGGAAGCTCATTT
GCAGACAAAGGCAAGTCCTTTTTCTTTCTCTTTTGGGCTTATCTCTGGCGGGCGGGCGGAA
CCTAGAAGCTATTCTGTGGTGGAGGAACTGAGGGCAGCTCCTTTGTACCAATTTAGCAAAGGA
CCTGGGTCTGGAGCAGAGGGAATTCTCCAGGCGGGGGTTAGGGTTGTTTCCAGAGGGAACAAAC
TACATTTGCAGCTCAATCAGGAGACCGCGGATTTGTTGCTAAATGAGAAATTGGACCGTGAGGAT
CTGTGCGGTCACACAGAGCCCTGTGTGCTACGTTTCCAAGTGTTGCTAGAGAGTCCCTTCGAGTT
TTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGACCACTCTCCAGTATTTCTGGACAAACAAA
TGTTGGTGAAAGTATCAGAGAGCAGTCTCTGGGACTACGTTTCTCTGAAGAATGCCGAAGAC
TTAGATGTAGGCCAAAACAATATTGAGAATAATATAATCAGCCCCAACTCCTATTTTCGGGTCT
CACCCGCAAACGCAGTGATGGCAGGAAATACCCAGAGCTGGTGCTGGACAAAGCGCTGGACCGAG
AGGAAGAAGCTGAGCTCAGGTAAACACTCACAGCACTGGATGGTGGCTCTCCGCCCAGATCTGGC
ACTGCTCAGGTCTACATCGAAGTCCTGGATGTCAACGATAATGCCCTGAATTTGAGCAGCCTTT
CTATAGAGTGCAGATCTCTGAGGACAGTCCGGTAGGCTTCTGGTTGTGAAGGTCTCTGCCACGG
ATGTAGACACAGGAGTCAACGGAGAGATTTCTATTCACTTTTCCAAGCTTCAGAAGAGATTGGC
AAAACCTTTAAGATCAATCCCTTGACAGGAGAAATTGAACTAAAAAAACAACCTCGATTTGAAAA
ACTTCAGTCCTATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTTCTGGAAAATGCACCG
TTCTGATTCAAGTGATAGATGTGAACGACCATGCCCCAGAAGTTACCATGTCTGCATTTACCAGC
CCAATACCTGAGAACGCGCCTGAAACTGTGGTTGCACTTTTCAGTGTTCAGATCTTGATTACAG
AGAAAATGGGAAAATTAGTTGCTCCATTCAGGAGGATCTACCCTTCCTCCTGAAATCCGCGGAAA
ACTTTTACACCCTACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACAACATCACT
ATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGCTGATCGC
CGATGTCAATGACAACGCTCCCGCCTTCACCCAAACCTCCTACACCCTGTTTCGTCCGCGAGAACA
ACAGCCCCCGCCTGCACATCCGCAGCGTCAGCGCTACAGACAGAGACTCAGGCACCAACGCCCAG
GTCACCTACTCGCTGCTGCCGCCCCAGGACCCGCACCTGCCCCCTCACATCCCTGGTCTCCATCAA
CGCGGACAACGGCCACCTGTTCCGCCCTCAGGTCTCTGGACTACGAGGCCCTGCAGGGGTTCCAGT
TCCGCGTGGGCGCTTCAGACCACGGCTCCCCGGCGCTGAGCAGCGAGGCCCTGGTGCGCGTGCTG
GTGCTGGACGCCAACGACAACCTCGCCCTTCGTGCTGTACCCGCTGCAGAACGGCTCCGCGCCCTG
CACCGAGCTGGTGCCCCGGGCGGCCGAGCCGGGCTACCTGGTGACCAAGGTGGTGGCGGTGGACG
GCGACTCGGGCCAGAACGCCTGGCTGTGCTACCAGCTGCTCAAGGCCACGGAGCTCGGTCTGTTC
GGCGTGTGGGCGCACAAATGGCGAGGTGCGCACCGCCAGGCTGCTGAGCGAGCGCGACGCGGCCAA
GCACAGGCTGGTGGTGTGGTCAAGGACAATGGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGC
ACGTGCTCCTGGTGGACGGCTTCTCCAGCCCTACCTGCCTCTCCCGGAGGCGGCCCCGACCCAG
GCCCAGGCCGACTTGCTCACCGTCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCCT
CTTTTCGGTGTCTCCTGTTTCGTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCCTCGGTGGGT
GCTGCTTGGTGCCCCAGGGCCCCCTTCCAGGGCATCTTGTGGACATGAGCGGCACCAAGGACCTA
TCCCAGAGCTACCAGTATGAGGTGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCT
GAAGCCGATTATCCCCAACTTCCCTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATTCTACCT
TCCCCAATAACTTTGGGTTCAATATTCAGTGACCATAGTTGACTTTTACATTCCATAGGTATTTT
ATTTTGTGGCATTTCATGCCAATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGA
TTTACTCTTGATTTTCTCATGTTCTTTCTCCCTTTGTTTTAAAGTGAACATTTACCTTTATTCC
TGTTCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSRRGV
RVVSRGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRVFQVLLESPFEFFQAELOVIDINDH
SPVFLDKQMLVKVSESSPPGTTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKRS DGRKYPEL
VLDKALDREEEAELRLTLTALDGGSPPRS GTAQVYIEVL DVNDNAPEFEQPFYRVQISEDSPVGF
LVVKVSATD VDTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDFEKLQSYEVNIEARDA
GTFSGKCTVLIQVIDVNDHAEVTMSAFTSPIPENAPETVVALFSVSDLDSENGKISCSIQEDL
PFLKSAENFYTLLTERPLDRESRAEYNITITVTDLGTPMLITQLNMTVLIADVNDNAPAFQTQS
YTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDPHLPLTSLVSINADNGHLFALRSLD
YEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNSPFVLYPLQNGSAPCTELVPRAAEPGYL
VTKVVAVDGDSGQNAWLSYQLLKATELG LFGVWAHNGEVRTARLLSERDAAKHRLVVLVKDN GEP
PRSATATLHVLLVDGFSQPYLPLPEAAPTQAQADLLTVYLVVALASVSSLFLFSVLLFVAVRLCR
RSRAASVGRCLVPEGPLPGHLVDMSGTRTLSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPG
KEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCGCGTAGCCGTGCGCC
GATTGCCTCTCGGCCTGGGCAATGGTCCCGGCTGCCGGTCGACGACCGCCCCGCGTCATGCGGCT
CCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTGGAGGTTGCAG
AGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCTCTCCAGGTGGGGGCTGTGTAC
CTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGCAGAAGAGGCCAATGC
GGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTGTGATTCTGGGGAAGCTG
AGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCTGGAGGAGCGGAGGACTCAAGG
TGCAACGTCCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGGAGCACACTTCCCTGACAGAGAAGA
GGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACGCAGCCCCGACAGAGGACTCCAATAACA
CTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAGGAGAGAAACATTACAGGATTAGAAAATTTCT
ACTCTGAAAATTTTAAATATGTACAGGACCTTATGGATTTTCTGAACCCAAACGGTAGTGA
TACTCTAGTCCTGTTTTACACCCCGTGGTGCCGCTTTTCTGCCAGTTTGGCCCCCTCACTTTAACT
CTCTGCCCCGGGCATTTCCAGCTCTTCACTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTT
TCTACCAGGTTTGGCACCGTAGCTGTTCTTAATATTTTATTATTTCAAGGAGCTAAACCAATGGC
CAGATTTAATCATAACAGATCGAACACTGGAAACACTGAAAATCTTCATTTTAAATCAGACAGGTA
TAGAAGCCAAGAAGAATGTGGTGGTAACTCAAGCCGACCAAATAGGCCCTCTTCCCAGCACTTTG
ATAAAAAGTGTGGACTGGTTGCTTGTATTTTCTTATTCTTTTAAATTAGTTTTATTATGTATGC
TACCATTGGAAGTGAAGTATTCGGTGGCTAATTCAGGACAAGAGCAGGAACATGTGGAGTAGT
GATGGTCTGAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTT
TCATACATTTTCTCCAGTGACGTGTTGACTTGAACTTCAGGCAGATTAAAGAATCATTTGTTG
AACAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGCA
AAAATATTCAATAG

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333

><subunit 1 of 1, 360 aa, 1 stop

><MW: 39885, pI: 4.79, NX(S/T): 7

MVPAAGRPRPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEEELL
HDPMGQDRAAEEANAVLGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRCNVRESL
FSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLENFTLKILNM
SQDLMDFLNPNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDASQHSSLSTRFGTV
AVPNILLFQGAQPMARFNHTDRTLETLETKIFIFNQGTGIEAKKNVVVTQADQIGPLPSTLIKSVDWL
LVFSLFFLISFIMYATIRTESIRWLIPGQEQEHVE

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCACGATGCTGCAGGGCCCTGGCTCGC
TGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGCGGGCTCTTCCTCTTTGGC
CAGCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCAGCTGTGCCA
CGGCATCGAATACCAGAACATGCGGCTGCCAACCTGCTGGGCCACGAGACCATGAAGGAGGTGC
TGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCGGACACCAAGAAGTTC
CTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGACCATCCAGCCATGCCACTC
GCTCTGCGTGCAAGTGAAGGACCGCTGCGCCCCGGTCATGTCCGCCTTCGGCTTCCCCTGGCCCG
ACATGCTTGAGTGCGACCGTTTCCCCCAGGACAACGACCTTTGCATCCCCCTCGCTAGCAGCGAC
CACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATGTGAAGCCTGCAAAAATAAAAATGATGA
TGACAACGACATAATGGAACGCTTTGTAAAAATGATTTTGCACTGAAAATAAAAGTGAAGGAGA
TAACCTACATCAACCGAGATACCAAAATCATCCTGGAGACCAAGAGCAAGACCATTTACAAGCTG
AACGGTGTGTCCGAAAGGGACCTGAAGAAATCGGTGCTGTGGCTCAAAGACAGCTTGCAAGTGCAC
CTGTGAGGAGATGAACGACATCAACGCGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGC
TGGTGATCACCTCGGTGAAGCGGTGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGC
ATCCGCAAGCTGCAGTGCTTAGTCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGG
CTGACCATTTCTGCTCCGGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCT
CAGCCTGGGCAGCTTCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTTCCTGAGTTATAAGG
CCACAGGAGTGGATAGCTGTTTTACCTAAAGGAAAAGCCCACCCGAATCTTGTAGAAATATTCA
AACTAATAAAATCATGAATATTTTAA

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920

><subunit 1 of 1, 295 aa, 1 stop

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPNLLG
HETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDRCAPVMS
AFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDNDNDIMETLCKNDFA
LKI KVKEITYINRDTKIILETKSKTIYKLNGVSEKDLKKSVLWLKDSLQCTCEEMNDINAPYLV
GQKQGGEVLVTSVKRWQKGQREFKRISRSIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCGACGATGGCGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGGCCC
TGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCTTGCTCGCGATCAGCCTG
CTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCAATTGCTGAGGA
CTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGGTATCCACCCCAT
TTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTGCGGGCTCCGTGCGGCGACCATCCTGGGTGCG
TGGCTGAACTTTGCCGGGAGTGTGCTACGCATGGTGGCCTGCATGGTTGTTGGGACCCAAAACCC
ATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCCAGAGCCTGGTCATCTTCTCTC
CAGCCAAGCTGGCTGCCTTGTGGTTCCCAGAGCACCAGCGAGCCACGGCCAACATGCTCGCCACC
ATGTCGAACCCTCTGGGCGTCCTTGTGGCCAATGTGCTGTCCCCTGTGCTGGTCAAGAAGGGTGA
GGACATTCCGTTAATGCTCGGTGTCTATACCATCCCTGCTGGCGTCGTCTGCCTGCTGTCCACCA
TCTGCCTGTGGGAGAGTGTGCCCCCACCCTGCTGCCGGGGCTGCCAGCTCCACCTCAGAG
AAGTTCCTGGATGGGCTCAAGCTGCAGCTCATGTGGAACAAGGCCTATGTCATCCTGGCTGTGTG
CTTGGGGGAATGATCGGGATCTCTGCCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAA
GCGGCCACTCCAGTGGGTTTTCCGGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGG
GCACTGGCTCTCGGCCCTATGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCT
GTGCCTGTTCTCTCTGGCCTGCGTGCCCTTTGCCCTGGTGTCCCAGCTGCAGGGACAGACCCTTG
CCCTGGCTGCCACCTGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAG
TTGGCGGTGAGTGTTCCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGG
GCAGGCCGAGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGT
CCTTGTCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCC
GGCCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCCTGCA
GGCCGAGTCTGGGGAGCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCGGGTG
TGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCTGGGGCCAGCACGGCGACTCCGGAGTGCACG
GCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCGGGAGCCCCACCAGCCTGCCACCGAGC
GACTCCCCGTGCGCAAGGCCCAGCAGCCACCGACGCGCCCTCCCGCCCCGGCAGACTCGCAGGCA
GGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTTCTCCTCCCCGTGGGTG
ATCACGTAGCTGAGCGCCTTGTAGTCCAGTTGCCCGCCACATCGATGGAGGCGAACTGGAACAT
CTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCTCCGGGAGCGAATTACAAGCGCG
CACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988

><subunit 1 of 1, 560 aa, 1 stop

><MW: 58427, pI: 6.86, NX(S/T): 2

MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLVLSM
EQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTQNPFAFLM
GGQSLCALAQSLVIFSPAKLAALWFPEHQ RATANMLATMSNPLGVLVANVLS PVLVKKGEDIPLM
LGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKAYVILAVCLGGM I
GISASFSALEEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTKHFTEATKIGLCFLSL
ACVPFALVSQ LQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPVGEGAATGMIFVLGQAEGI
LIMLAMTALT VRRSEPSLSTCQQGEDPLDWTVSLLL MAGLCTFFSCILAVFFHTPYRR LQAESGE
PPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARGASLEDPRGPGSPHPACHRATPRAQ
GPAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 170

GTCCACATCCTGCTCAACTGGGTCAGGTCCCTCTTAGACCAGCTCTTGTCCATCATTGCTGAAGTGGACCAACTAG
TTCCCCAGTAGGGGGTCTCCCCTGGAATCTTGTATCGGCGTTTGGACATCTCAGATCGCTTCCAATGAAGATGGCCT
TGCCTTGGGGTCTGCTTGTGTTTATAATCATCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGGAAGGAGCAC
GGGGCTGATCAAGCCATCCAGGAAACTGAGGACTTGTCCAGCCTTGAAAGAACTCTAGTGGTTTCTGAATCTAGC
CCACTTGGCGGTAAGCATGATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGGTGGCTACTTATTTCTTTT
AGGGGATTGTGAGGAGGTGACCACTCTCACGGTGAAATACCAAGTGTGAGAGGAAGTGCCATCTGGTACAGTGATCGG
GAAGCTGTCCCAGGAACGGGGCCGGGAGGAGGGCGGAGGCAAGCTGGGGCCGCTTCCAGGTGTTGCAGCTGCCTCA
GGCGCTCCCCATTACAGGTGGACTCTGAGGAAGGCTTGTCTAGCACAGGCAGGCGGCTGGATCGAGAGCAGCTGTGCCG
ACAGTGGGATCCCTGCCTGGTTTCTTGTATGTGCTTGGCACAGGGGATTGGCTCTGATCCATGTGGAGATCCAAGT
GCTGGACATCAATGACCACCAGCCACGGTTTCCCAAAGGCGAGCAGGAGCTGGAAATCTCTGAGAGCGCCTCTCTGCG
AACCCGGATCCCCCTGGACAGAGCTCTTGACCCAGACACAGGCCCTAACACCCTGCACACCTACACTCTGTCTCCCAG
TGAGCACTTTGCCCTTGGATGTCAATGTGGGCCCTGATGAGACCAACATGCAGAACTCATAGTGGTGAAGGAGCTGGA
CAGGGAATTCATTCATTTTGTATCTGGTTAACTGCCCTATGACAATGGGAACCCCCCAAGTCAGGTACCAGCTT
GGTCAAGGTCAACGTCTTGGACTCCAATGACAATAGCCCTGCGTTTGTGAGAGTTCACTGGCACTGGAATCCAAGA
AGATGCTGCACCTGGTACGCTTCTCATAAACTGACCGCCACAGACCTGACCAAGGCCCCAATGGGGAGGTGGAGTT
CTTCCTCAGTAAGCACATGCCTCCAGAGGTGCTGGACACCTTCACTATTGATGCCAAGACAGGCCAGGTCACTTCTGCG
TCGACCTCTAGACTATGAAAAGAACCTGCTTACGAGGTGGATGTTTCAAGCAAGGGACCTGGGTCCCAATCCTATCCC
AGCCATTGCAAAGTTCTCATCAAGGTTCTGGATGTCAATGACAACATCCCAAGCATCCACGTACATGGGCCCTCCCA
GCCATCACTGGTGTGAGAAGCTCTTCCAAGGACAGTTTTATTGCTCTTGTATGGCAGATGACTTGGATTCAAGGACA
CAATGGTTTGGTCCACTGCTGGCTGAGCCAAGAGCTGGGGCACTTCAGGCTGAAAAGAACTAATGGCAACACATACAT
GTTGCTAACCAATGCCACACTGGACAGAGAGCAGTGGCCCAATATACCCTCACTCTGTTAGCCCAAGACCAAGGACT
CCAGCCCTTATCAGCCAAAGAAACAGCTCAGCATTAGATCAGTGCATCAACGACAATGCACCTGTGTTTGAAGAAAG
CAGGTATGAAGTCTCCACGCGGGAACAACTTACCCTCTCTTACCTCATTACCATCAAGGCTCATGATGCAGACTT
GGGCATTAATGGAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGG
AGAGGTCACTGCTCAGAGGTCACTGAATATGAAGAGATGGCCGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG
GCAACCCATTGCTGCATCCAGTGTCTGTGTGGTGCAGCCTTGGATGCCAATGATAATGCCCCAGAGGTGGTCCA
GCCTGTGCTCAGCGATGGAAAGCCAGCCTCTCCGTGCTTGTGAATGCCCTCCACAGGCCACCTGCTGGTGGCCATCGA
GACTCCCAATGGCTTGGGCCCAGCGGCCACTGACACACCTCCACTGGCCACTCACAGCTCCCGGCCATTCTTTTGAC
AACCATTGTGGCAAGAGATGCAGACTCGGGGGCAAATGGAGAGCCCTCTACAGCATCCGCAATGGAAATGAAGCCCA
CCTCTTCATCCTCAACCCTCATACGGGGCAGCTGTTCCGTCAATGTACCAATGCCAGCAGCCTCATTGGGAGTGAGTG
GGAGCTGGAGATAGTAGTAGAGGACCAGGGAAGCCCCCTTACAGACCCGAGCCCTGTTGAGGGTCATGTTTGTAC
CAGTGTGGACCACCTGAGGGACTCAGCCCGCAAGCCTGGGGCCTTGGATGTGATGCTGACGGTGTCTGCTGGC
TGTACTGTTGGGCATCTTCGGGTGATCCTGGCTTTGTTTCAATGTCCATCTGCCGGACAGAAAAGAGGACAACAGGGC
CTACAACGTGTCGGGAGGCCGAGTCCACCTACCGCCAGCAGCCCAAGAGGCCCCAGAAAACATTGAGAAGGCAGACAT
CCACCTGCTGCCTGTGCTCAGGGTCAAGGAGTGCAGCCTTGTGAAGTCCGGCAGTCCCAAAAGATGTGGACAAGAGCT
GGCGATGATGGAAGCAGGCTGGGACCCCTGCCTGCAGGCCCTTCCACCTCACCCGACCTGTACAGGACGCTGCG
TAATCAAGGCAACCAGGGAGCACCGGCGGAGAGCCGAGAGGTGCTGCAAGACACGGTCAACCTCCTTTTCAACCATCC
CAGGCAGAGGAATGCCCTCCCGGAGAACCTGAACCTTCCCGAGCCCAAGCCTGCCACAGGCCAGCCACGTTCCAGGCC
TCTGAAGGTTGCAGGCAGCCCAAGGGAGGCTGGCTGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCACCAGC
CTCCTCTGCAACCCCTGAGACGGCAGCGACATCTCAATGGCAAAGTGTCCCTGAGAAAGAATCAGGGCCCCGTCAGAT
CCTGCGGAGCCTGGTCCGGCTGTCTGTGGCTGCCTTCCCGAGCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCC
TCCTGTTTCAAGCAATCTCCAGCTGCTGTCTTGTGTCATCAGGGCCAATTCAGCCCCAAACCAACACCGAGGAAA
TAAGTACTTGGCCAAGCCAGGAGGCAGCAGGAGTGCAATCCAGACACAGATGGCCCAAGTGCAAGGGCTGGAGGCCA
GACAGCCCCAGAACAGGAGGAAGGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGAAGCACTGCTAGAAGAAGAGCT
GTCAAGTCTGCTGGACCCAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCCCTGACCCGGCCTGGATGGCGAGACT
CTCTTTGCCCTTACCACCAACTACCGTGACAATGTGATCTCCCGGATGCTGCAGCCACGGAGGAGCCGAGGACCTT
CCAGACGTTTCGGCAAGGCAGAGGCACCAGAGCTGAGCCCAACAGGCACGAGGCTGGCCAGCACCTTTGTCTCGGAGAT
GAGCTCACTGCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCGTGGAGGCCGCTCCGAGGCGCTGCGGCGGCT
CTCGGTCTGCGGGAGGACCTCAGTTTAGACTTGGCCACCAGTGCAGCCTCAGGCATGAAAGTGCAAGGGGACCCAGG
TGAAAGACCGGGACTGAGGGCAAGAGCAGAGGCAGCAGCAGCAGCAGGTCCTGTGAACATACCTCAGACGCCT
CTGGATCCAAGAACCAGGGCCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAAATCTTGTAACTCACTAGCTAGCGG
CGGCCTGAGAACTTTAGGGTACTGATGCTACCCCCACAGAGGAGGCAAGAGCCCCAGGACTAACAGCTGACTGACCA
AAGCAGCCCTTGTAAAGCAGCTCTGAGTCTTTTGGAGGACAGGGACGGTTTGTGGCTGAGATAAGTGTTCCTGGCAA
AACATATGTGGAGCACAAAGGCTCAGTCTCTGGCAGAACAGATGCCACGAGTATCAAGGCAGGAAAGGGTGGCCT
TCTTGGGTAGCAGGAGTCAGGGGGCTGTACCCTGGGGGTGCCAGGAAATGCTCTCTGACCTATCAATAAAGGAAAAGC
AGTAAAAAAAAAAAAAAAAAAAAA

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLLGLLGGPYLFLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQAGAA
FQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQVLDIND
HQPFRPKGEQELEISESASLRTRIPDLRALDPDTGPNTLHTYTLSPSEHFALDVIVGPDETKHAE
LIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSDNDNSPAFAESSLALEIQEDAAPG
TLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPLDYEKNPAYEVDVQAR
DLGPNPIPAHCKVLIKVLVDVNDNIPSIHVTWASQPSLVSEALPKDSFIALVMADDLDSGHNGLVH
CWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQDQGLQPLSAKKQLSIQISDIND
NAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVSYRIQDSPVAHLVAIDSNTGEVTAQ
RSLNYEEMAGFEFQVIAEDSGQPMLASSVSVVWSLLDANDNAPEVVQPVLSDGKASLSVLVNA
GHLLVPIETPNGLGPAGTDTPLATHSSRPFLTTIVARDADSGANGEPLYSIRNGNEAHLFILN
PHTGQLFVNVTNASSLIGSEWELEIVVEDQGSPPLOTRALLRVMFVTSVDHLRDSARKPGALSMS
MLTVICLAVLLGIFGLILALFMSICRTEKKDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPV
LRGQAGEPCEVGQSHKDVDKEAMMEAGWDPCLQAPFHLTPTLYRTLNRNQGNGAPAESREVLQDT
VNLLFNHPRQRNASRENLLPEPQPATGQPRSRPLKVAGSPTGRLAGDQGSSEAPQRPASSATL
RRQRHLNGKVSPEKESGPRQILRSLVRLSVAFAERNPVEELTVDSPPVQQISQLLSLLHQGFQ
PKPNHRGNKYLAKEGSRSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLD
PSTGLALDRLSAPDPAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRL
ASTFVSEMSSLLEMLLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAAAGMKVQGDPPGKTGT
EGKSRGSSSSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAG
ACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGA
GCCTTCCTTACACTTCGCCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATAC
TATTTTTTGGATTGTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAG
TATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGCACCATGTTTGAGCTCATCAT
CTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTG
TAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATC
CGACTACTGCATAAACAACGACTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTT
CTGGAAACTAGGAGATCCCTTTCCCATTTCTCAGCCCAAAACATGGGATCTTATCCATAGAACAGC
TCATCAGCCGGGTTGGTGTGATTGGAGTGAAGTCTCATGGCTCTTCTTTCTGGATTGGTGTCTGTC
AACTGCCCATACACTTACATGTCTTACTTCCTCAGGAATGTGACTGACACGGATATTCTAGCCCT
GGAACGGCGACTGCTGCAAACCATGGATATGATCATAAGCAAAAAGAAAAGGATGGCAATGGCAC
GGAGAACAATGTTCCAGAAGGGGGGAAGTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAA
AGTGTTACCACTTCAGCATCAGGAAGTGAAGTCTTACTCTTATTCAACAGGAAGTGGATGCTTT
GGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAG
AATACTCCAAAACCTTCAAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGT
GTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGT
CACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCC
AACACATTTCTTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACT
CTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCATTGTCTCTGCTATTAGC
ACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAAT
ACCGCACCATAATCACTGAAGTCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTGAT
GTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACC
AGAGAAGCAAATGGCACCTTGAAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTTCAAAA
TTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAACAACAAAATGCTA
TGGTAGCATTTTTACCTTCATAGCATACTCCTTCCCCGTCAGGTGATACTATGACCATGAGTAG
CATCAGCCAGAACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTG
GATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAAC
CTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTACAT
GGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATTAAAATCAGAGACTGTAACAAAAA
AAAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATTAAAATCAGAGACTGTAACAAAAA
AAAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATTAAAATCAGAGACTGTAACAAAAA
TTGTTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFASFCTMFELIIFEILGV
LNSSSRYPFHWMNLCVILLILVFMVPPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGDP
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDAL EELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDVPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKEFFYAISSSKSSNVIVLLLAQIMGY
FVSSVLLIRMSMPLEYRTIITEVLGELQNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCATCA
TGATTACCTCCCNGANACTATTTTTTGGATTTGGGTGGCTTTTCTTCNGCGCCAATGTTTAAAGA
CTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTTCTTGACCA
TGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGTTATTTTCACTGG
AAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTTTTTACATTGGCTATTT
TATTGTGAGCAATATCCGACTACTGCATAACAACGACTGCTTTTTTCCTGTCTCTTATGGCTGA
CCTTTATGTATTTCCAG

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAATTTGGCCAAGGGTTTCTTT
NTTGAATTCGGGTTNNGNATACCTTCCCAGAAAATATTTTTTGGATTTGGGGTAGNTTTTTTTC
ATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTTNTCCGTGAC
GTTTGCATTTTCTTGCACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAGTATTGAATAGCA
GCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTTCATGGTG
CCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTT
TTCCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAANTAGGAGATCCCTTCCCATTCTC
TC

FIGURE 176

CTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCCGCGACCCTTGGGGGGCCTCCGGGATTGCTACCTTTTTGG
CTCCCTGCTCGTGAAGTCTCTTCTACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGGCGAAGGAGGG
CGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAAGTTGCAGCCCCGACCCAGAGCTGGCTGTGGT
GGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCGAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTGGCCGTTGAGCCT
GGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACCAGTGGTT
GGGAGTCAGTGTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCGATATGAGGCAAGGCAGCGAGT
GGACCATCTTGGAGACGCGGGATATGATTGGTCTGCTCTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTT
GGATGGTGGGAAAGTCTTCATCTACCATGGGAGCAGCCTGGGGGTTGTGCGCAAACCTTACAGGTGCTGGAGGGCGA
TGCCGCTTCTCCCTGATAGCCACTACCTCCTCTTGGGGCCCCAGGAACCTATAATTGGAAGGGCACGGCCAGGGT
GGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACGACGCTCCCTACGAGGCGGGGGGAGAGAAGGAGCA
GGACCCCGCCTCATCCCGTCCCTGCCAACAGCTACTTGGCTTCTCTATTGACTCGGGGAAAGGTCTGGTGCCTGC
AGAAGAGCTGAGCTTGTGGCTGGAGCCCCCGCGCCAACCACAAGGGTGTGTGGTTCATCCTGCGCAAGGACAGCGC
CAGTCGCTTGGTGCCGAGGTTATGCTGTCTGGGAGCGCCTGACCTCCGGCTTGGCTACTCACTGGCTGTGGCTGA
CCTCAACAGTGTGGCTGGCCAGACCTGATAGTGGGTGCCCCCTACTTCTTGGAGCGCCAAGAAGAGCTGGGGGGTGC
TGTGTATGTGACTTGAACAGGGGGGTCACTGGGCTGGGATCTCCCTCTCCGGCTCTGCGGCTCCCTGACTCCAT
GTTCCGGATCAGCCTGGCTGTCTGGGGGACCTCAACCAAGATGGCTTTCAGATATTGCAGTGGGTGCCCCCTTGA
TGGTGTGGGAAAGTCTTCATCTACCATGGGAGCAGCCTGGGGGTTGTGCGCAAACCTTACAGGTGCTGGAGGGCGA
GGCTGTGGGCATCAAGAGCTTCGGCTACTCCCTGTGAGGAGCTTGGATATGGATGGGAACCAATACCCTGACCTGCT
GGTGGGCTCCCTGGCTGACACCGCAGTGTCTTCAGGGCCAGACCATCCTCCATGTCTCCCATGAGGTCTCTATTGC
TCCACGAAGCATCGACCTGGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAG
CTACATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCCGCTTCTGTACCCGGGTGAGCGACACAGACCCGGAGCT
CCGGGGCCAGGTTCCCCGTGTGACGTTCTGAGCGCTAACCTGGAAGAACCAAGCACCAGGCCTCGGGCACCGTGTG
GCTGAAGCACCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGC
CATTGTAGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGCTGCCTCCAGT
GGCCCCATCCTCAATGCCACCAGCCAGCACCAGCGGGCAGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGA
CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCCGCTTCTGTACCCGGGTGAGCGACACAGGAATTCCAACCTCT
GCCCCATGGATGTGGATGGAACAACAGCCCTGTTTGCAGTGTGGGAGCCAGTCAATTGGCCTGGAGCTGATGGTCAC
CAACCTGCCATCGGACCCAGCCAGCCAGCCAGGCTGATGGGGATGATGCCCATGAAGCCAGCTCCTGGTCACTGCTTCC
TGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCCTGCGGAGAAGCCACTCTGCCTGTCCAATGAGAATGCCCTC
CCATGTTGAGTGTGAGTGGGGAAACCCATGAAGAGAGGTGCCCCAGGTACCTTCTACCTCATCCTTAGCACCTCCGG
GATCAGCATTGAGCAACCGAAGTGGAGGTGAGAGTGTGTTGCCACAGATCAGTGAGCAGGAGCTGCATCCAGTCTC
TGCACGAGCCCGTGTCTTCAATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCCCCAGCAACTCTTCTTCTCTGG
TGTGGTGTGGGGCGAGAGAGCCATGCAGTCTGAGCGGGATGTGGGCGCAAGGTCAAGTATGAGGTACCGTTTCCAA
CCAAGGCCAGTCGCTCAGAACCTGGGCTCTGCCTTCTCAACATCATGTGGCCTCATGAGATTGCCAATGGGAAGTG
GTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCTGGGCGAGAAAGGGCTTGTCTCTCCAGGCCCAA
CATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGGGAGCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGA
GCGGCAGGAGCCAGCATGTCTGGTGGCCAGTGTCTCTGCTGAGAAGAAGAAAAACATCACCTGGACTGCGCCCG
GGGCACGGCCAACCTGTGTGGTGTTCAGCTGCCCACTCTACAGCTTGTACCGCGCGGCTGTGCTGCATGTCTGGGGCCG
TCTCTGGAACAGCACCTTCTGGAGGAGTACTCAGCTGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCAGT
GAAGTCTCCATAAAGAACTTGATGCTCCGAGATGCCTCCACAGTGATCCCACTGATGGTATACTTGGACCCCATGGC
TGTGGTGGCAGAAGGAGTGCCTGGTGGGTGATCCTCCTGGCTGTACTGGCTGGGCTGCTGGTGTGCTAGCACTGCTGGT
GCTGCTCCTGTGAAGATGGGATTCTTCAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAA
GATTCTCGGGAAGACCGACAGCAGTTCAGGAGGAGAAGACGGGCACCATCCTGAGGAACAACCTGGGGCAGCCCCCG
GCGGAGGGGCGGATGCACACCCCATCCTGGCTGTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCC
AGGCACCGCTAGGTTCCCATGTCCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGA
TGAAGAGGGTAGAGTGGGCTGCTGGTGTGCGATCAAGATTTGGCAGGATCGGCTTCCCTCAGGGGCACAGACCTCTCCC
ACCCACAAGAACTCCTCCCACCCAACTCCCCCTTAGAGTGTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCAT
GGGGTAGGGTGAGAAGGGCAGGGGTGTCTGATGCAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCATTCA
CCCTGTGTAACAGGACCCCAAGGACCTGCCTCCCCGGAAGTGCCTTAACCTAGAGGGTCGGGGAGGAGGTTGTGTAC
TGAATCAGGCTGCTCCTTCTCTAGTTTCCCCTCTCATCTGACCTTAGTTTGTGCTGCCATCAGTCTAGTGGTTTCGTGGT
TTCGTCTATTTATTAATAAATATTTGAGAACAATAAAAAAAAAAAAAAAAAAAAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQLQPR
PQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWLGVSVRS
QGGPGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGGEWKFCGRPQGHEQFG
FCQQGTAAAFSPDSHYLLFGAPGTYNWKG TARVELCAQGSADLAHLDDGPYEAGGEKEQDPRILP
VPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLVPEVMLS GERLTSGFG
YSLAVADLNSDGWPD LIVGAPYFFERQEELGGAVYVYLNQGGHWAGISPLRLCGSPDSMFGISLA
VLGDLNQGDFPDIAVGAPFDGDGKVF IYHGSSLGVVAKPSQVLEGEAVGIKSFGYSLSGSLDMDG
NQYPDLLVGSLADTAVLFRARPILHVSHEVSIAPRSIDLEQPN CAGGHSVCVDLRVCFSYIAVPS
SYSPTVALDYVLDA DTDRRLRGQVPRVTFLSRNLEEPKHQASGT VWLKHQHDRVCGDAMFQLQEN
VKDKLRAIVVTLSSYSLQTPRLRRQAPGQGLPPVAPILNAHQ PSTQRAEIHFLKQCGEDKICQSN
LQLVHARFCTRVSDTEFQPLPMDVDGTTALFALSGQPVIGLELMVTNLPSDPAQPQADGDDAHEA
QLLVMLPDSLHYSGVRALDPAEKPLCLSNENASHVECELGNPMKRGAQVTFYLILSTSGIS IETT
ELEVELLLATISEQELHPVSARARVFIELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVK
YEVTVSNQGSRLTGLSAFLNIMWPHEIANGKWLLYPMQVELEGGQGPQKGLCSPRPNILHLDV
DSRDRRRRELEPPEQQEPGERQEPSMSWWPVSSAEKKKNITLDCARGTANCVVFS CPLYSFDRAA
VLHVWGRLWNSTFLEEYS AVKSLEIVRANITVKSSIKNLMLRDASTVIPVMVYLDPM AVVAEGV
PWVVILLAVLAGLLVLALLVLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILR
NNWGSPRREGPD AHPILAADGHPELGPDPGHPGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-
407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408
and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGAAAG
CAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCTCACAAAC
AAGATGCTCAAGGTGTCAGCCGTA CTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCAGTCTCTCGC
AGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTTCGGACGGCGGTAATTTTCTGGATGATAAAC
AATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAACAAATTCCGAGACGAA
GTAGAGGATGATTATTTCCGCACCTTGGAGTCCAGGAAAACCTTCGATCAGGCTTTAGATCCAGC
TAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTATGCATTGCTCAAGATTCTCAGA
CTGCAGTCTGCATTAGTCAACGGAGGCTTACACACAGGATGAAAGAAGCAGGAGTAGACCATAGG
CAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTGCCAGTGGTCTATCCCAGCCCTGTTTG
TGGTTCAGATGGTCATACCTACTCTTTTTCAGTGCAAAC TAGAATATCAGGCATGTGTCTTAGGAA
AACAGATCTCAGTCAAATGTGAAGGACATTGCCCATGTCTTTCAGATAAGCCCACCACTACAAGC
AGAAATGTTAAGAGAGCATGCAGTGACCTGGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTG
GTTCAAGGCCCTTCATGAAAGTGGAAGTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGA
GAAGCAGATTCGATACCAGCATCTTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAACAGA
CTTGATACAAACTATGACCTGCTATTGGACCAGTCAGAGCTCAGAAGCATTACCTTGATAAGAA
TGAACAGTGTACCAAGGCATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATA
ATGAGTGGTGCTACTGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATT
CAGAAGCGGCAAGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTA
CTACAAGCCAACACAATGTCATGGCAGTGTTGGACAGTGCTGGTGTGTTGACAGATATGGAAATG
AAGTCATGGGATCCAGAATAAATGGTGTGTCAGATTGTGCTATAGATTTTGAGATCTCCGGAGAT
TTTGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATGA
TGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGTGATGACCATG
ATGTATACATT**TG**ATTGATGACAGTTGAAATCAATAAATTCTACATTTCTAATATTTACAAAAAT
GATAGCCTATTTAAATTTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCACATATATTTT
GTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACCTTTATGTTTAAATAAGAATCATTGTC
TTTGAGTTTTTATATTCCTTACAAAAAGAAAATACATATGCAGTCTAGTCAGACAAAATAAAG
TTTTGAAGTGCTACTATAATAAATTTTTTACGAGAACAACTTTGTAAATCTTCATAAGCAAAA
TGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAGATAATTCTAAGTGAAATTTAAA
ATAAATAAATTTTAAATGACCTGGGTCTTAAGGATTTAGGAAAAATATGCATGCTTTAATTGCAT
TTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAGGATAACAGAGAGATACCACATGACTCCA
AAAAAAAAAAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829

><subunit 1 of 1, 436 aa, 1 stop

><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFRDEV
EDDYFRTWSPGKPFQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEAGVDHRQ
WRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPGPSDKPTSTSR
NVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLRPERSRFDTSILPICKDSLGMFNR
DTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQRQDPPCQTELSNIQ
KRQGVKKLLGQYIPLCDEDGYKPTQCHGSVGQCWCVDTRYGNEVMGSRINGVADCAIDFEISGDF
ASGDFHEWTDDEDDDDIMNDEDEIEDDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGCCTC
TTTCAGCCCGGATCGCCCCAGCAGGGATGGGCGACAAGATCTGGCTGCCCTTCCCGTGCTCCTTCTGGCCGCTCTG
CCTCCGGTGTCTGCTGCCTGGGGCGGCCGGCTTCACACCTTCCCTCGATAGCGACTTCACCTTTACCTTCCCGCCGGC
CAGAAGGAGTGCTTCTACCAGCCCATGCCCTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTTAGATGGAGCAGGA
TTAGATATTGATTTCCATCTTGCTCTCCAGAAGGCAAAACCTTAGTTTTGAACAAAGAAAATCAGATGGAGTTCAC
ACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTTCAGCACCATTCTGAGAAGGTGATTTTC
TTTGAATTAATCTGGATAATATGGGAGAACAGGCACAAGAACAAGATTGGAAGAAATATATTACTGGCACAGAT
ATATTGGATATGAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCCAGACTAAGCAAAAGTGGGCACATA
CAAATCTGCTTAGAGCATTGGAAGCTCGTGATCGAAACATACAAGAAAGCAACTTTGATAGAGTCAATTTCTGGTCT
ATGGTTAATTTAGTGGTCATGGTGGTGGTGTGAGCCATTCAAGTTTATATGCTGAAGAGTCTGTTTGAAGATAAGAGG
AAAAGTAGAACTTAAACTCCAACTAGAGTACGTAACATTGAAAAATGAGGCATAAAAAATGCAATAAACTGTTACAG
TCAAGACCATTAAATGGTCTTCTCCAAAATATTTTGAGATATAAAAGTAGGAAACAGGTATAATTTAATGTGAAAAT
AAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAGTTGTACTTAAGTGTGTAACAGGAATATTTTGAGAATAT
AGGTTTAACTGAATGAAGCCATATTAATACTGCATTTTCTTAACCTTTGAAAAATTTTGCAAATGTCTTAGGTGATTT
AAATAAATGAGTATTGGGCCCTAATTGCAACACCAGTCTGTTTTTAACAGGTTCTATTACCCAGAAGCTTTTTTGTAAT
GCGGCAGTTACAAATTAAGTGTGGAAGTTTTAGTTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGA
ATAAATCTTTAGACTACAAAAGCCCACTTTTCTCTATTTACATATGCATCTCTCTATAATGTAAATAGAATAATAG
CTTTGAAATACAATTAGGTTTTTGAGATTTTTATAACCAATACATTTTCAGTGTAACATATTAGCAGAAAGCATTAGT
CTTTGTACTTTTGCTTACATTTCCAAAAGCTGACATTTTACAGATTCTTAAAAACACAAAGTTACACTTACTAAAATTA
GGACATGTTTTCTCTTTGAAATGAAGAATATAGTTTAAAAGCTTCTCTCCATAGGGACACATTTTCTCTAACCTT
AATAAAGTGTAGGATTTTAAATTAATGTGAGGTAAAATAAGTTTATTTTAAATAGTATCTGTCAAGTTAATATCT
GTCAACAGTTAATAATCATGTTATGTTAATTTTAACATGATTGCTGACTTGGATAATTCATTATTACCAGCAGTTATG
AAGGAAATATTGCTAAAATGATCTGGGCCTACCATAAATAAATATCTCTTTTCTGAGCTCTAAGAATTATCAGAAAA
CAGGAAAGAATTTAGAAAACTTGAGAAAACCTAATCCAAAATAAAATTCACCTTAAGTAGAAGTATAAATAAATATCT
AGAATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
AACTTTGGCTGTAGGTTTTATTTTCTACAAGAATTCTGGTTTGAATTTATTTTGTAAAGCAGGTACATTTTATAAAA
TGTAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTATTAATAAATTTTATTATAACAATTTTATTAAAT
GGCCTTTCTGAACATTTATTTATTTATGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTAAACACCTAAT
GTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTTGAAAGTCTATGGGGGTCTTACTCAAGTACTAGT
AATTTAATTCATCATGAATGAAGTATAATTTTAAAGTTATGCCATTATAACGTTGTTTATGACTACATTGTGAGT
TAGAAACAACTTAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATTTCTGATGAGCAATAATGA
TAACCAGAGAGTGATTTTCACTCATAGTAGTATAAAAAGAGATACATTTCCCTCTTAGGCCCCCTGGGAGAAGA
GCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAATGCCGTATATGATCAATTACCTTAATTGGCCA
AGAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAAGCAAAAGGTCAATAAGATCCTTGCCTATGAAT
ACCCCTCCCTTTTGGCTGTATAAATTTGCAATGAGAAGCAAATTTACAGTACCATAACTAATAAAGCAGGGTACAGAT
ATAAACTACTGCATCTTTCTATAAACTGTGATTAAGAATTCTACCTCTCTGTATGGCTGTTACTGTACTGTACTC
TCTGACTCCTTACCTAACAAATGAATTTGTTACATAATCTTCTACATGTATGATTTGTGCCACTGATCTTAAACCTATG
ATTAGTAACCTCTTACCATATAAAAACGATAATTGCTTTATTTGGAAAAGAATTTAGGAATACTAAGGACAATTATT
TTTATAGACAAAGTAAAAAGACAGATATTTAAGAGGCATAACCAAAAAAGCAAACTTGTAACAGAGTAAAAATCTT
TAATATTTCTAAAGACATACTGTTTATCTGCTTCATATGCTTTTTTTAATTTCACTATTCCATTCTAAATTAAGTT
ATGCTAAATTGAGTAAGCTGTTTATCACTTAACAGCTCATTTTGTCTTTTCAATATACAAATTTTAAAAATACTACA
ATATTTAACTAAGGCCCAACCGATTTCCATAATGTAGCAGTTACCGTGTTCACCTCACACTAAGGCCTAGAGTTTGCT
CTGATATGCATTTGGATGATTAATGTTATGCTGTTCTTTATGTGAATGTCAAGACATGGAGGGTGTGTTGTAATTTTA
TGGTAAATTAATCCTTCTTACACATAATGGTGTCTTAAATTTGACAAAAATGAGCACTTACAATTTGATGTCTCCT
CAAATGAAGATTCTTTATGTGAAATTTTAAAGACATTGATTCCGCATGTAAGGATTTTTTCTGAGTACAATAAT
GCACAATCAGTGTGCTCAAAGCTTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATA
TGTATATAATAAAATTTATCAAAGGAAAA

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196

><subunit 1 of 1, 229 aa, 1 stop

><MW: 26017, pI: 4.73, NX(S/T): 0

MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLDSDFTFTLPAGQKECFYQPMPLKASLEIEYQVL
DGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFFELILDNMG
EQAQEQEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRNIQESNFDRV
NFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATAACCAGATCTC
ACCAGAGAGTCGCAGACACTATGCTGCCTCCCATGGCCCTGCCAGTGTGTCCTGGATGCTGCTT
TCCTGCCTCATTCTCCTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAAGTGCCTCTCCACG
GATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGTTTTTGTACCAA
AATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAAGTGGTGTCTGTGCTC
AGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAGTAACAGCTACTCATACAT
CTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATGGAGATGGATGGGAGTGGAGTA
GCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCCTCCACCATCTTAAACCCTGGCCAC
TGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTGGAAAGATTATAACTGTGATGCAAAGTT
ACCCTATGTCTGCAAGTTCAAGGACTTAGGGCAGGTGGGAAGTCAGCAGCCTCAGCTTGCGTGCA
GCTCATCATGGACATGAGACCAGTGTGAAGACTCACCCCTGGAAGAGAATATTCTCCCCAAACTGC
CCTACCTGACTACCTTGTCATGATCCTCCTTCTTTTCTTTTCTTCACCTTCATTTCAGGCTT
TTCTCTGTCTTCCATGTCTTGAGATCTCAGAGAATAATAATAAAAATGTTACTTTATAAAAAAA
AAAAAAAAAAAAAA

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

<subunit 1 of 1, 175 aa, 1 stop

<MW: 19330, pI: 7.25, NX(S/T): 1

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWMDAD
LACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSSTDVMNY
FAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

FIGURE 184

CCAGTCTGTGCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGCACA
GAGGAGTGGGCCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGCTGGCTGCCTG
CGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGTCCGACTGTG
TCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTACTCCCGGGAGATA
GTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAAGTGTAAGCCCTCGGA
TGTGGATGGCATCGGCCAGACCCTGCCCCTGTCTGCTGCAATACTGAGCTGTGCAATGTAGACG
GGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTCCTCCCACTCTTGAGCCTCCGA
CTGTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCAGCCCCGAATGCCTTGAAGAAGTGC
CCCCTGCACCAGGAAAAAAAAAAAAAAAAA

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYPFQG

DSTVTKSCASKCKPSDVGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSLRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGCACG
GTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTCTCCCTCTTGAGTCC
TTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGGTAGCGGCGG
CTCTCGGCGGCCACCTCTGCTGGGAGTGAGCGCCACCTTGAACCTCGGTTCTCAATTCCAACGCT
ATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTCTGCAGTCAGCGCCGC
GCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACAACCTACCAGCCGTACCCGT
GCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGTCCCACCCGCGGAGGGGACGCA
GGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACGCTGCATGCGTCACGCTATGTGCTG
CCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTTCTGATCAAAATCATTTCCGAGGAGAAA
TTGAGGAAACCATCACTGAAAGCTTTGGTAATGATCATAGCACCTTGGATGGGTATTCCAGAAGA
ACCACCTTGTCTTCAAAAATGTATCACACCAAAGGACAAGAAGGTTCTGTTTGTCTCCGGTCATC
AGACTGTGCCTCAGGATTGTGTTGTGCTAGACACTTCTGGTCCAAGATCTGTAAACCTGTCCTGA
AAGAAGGTCAAGTGTGTACCAAGCATAGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGT
TGTTACTGTGGAGAAGGTCTGTCTTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTC
TAGGCTTCACACTTGTGAGAGACACTAAACCAGCTATCCAAATGCAGTGAACCTCTTTATATAA
TAGATGCTATGAAAACCTTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGG
TTTCAGTTAAGCATTCCAATAACACCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATG
GAACTCCCCTGTGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAA
TGCAATGAACTTTTAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTTCTCTTGTATGTA
AATTTTTGTACACATTGATTGTTATCTTGACTGACAAATATTTCTATATTGAACTGAAGTAAATCA
TTTCAGCTTATAGTTCTTAAAAGCATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCAA
GGATCTCTTGGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATTTTCT
GAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGAAATAAAA
TTTAACATTTAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSAAPG
ILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRKRRCMRHAMCCPG
NYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEGSVCLRSSDC
ASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQKDHHQASNSSRL
HTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTTCCCTGCAGTCAGAAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGAACG
GGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTTCTTTCTCCTTCNKGAGTCC
TTNTGAGANGATGGTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATGGTAGCGGCG
GTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTCAATTCCAACGNT
ATCAAGAACCTGCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTTTCAGTCAGCGCCGC
GCCGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACAATTACCAGCCGTACCCGT
GCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGTCCCACCCGCGGAGGGGANGCG
GGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACGCTGCATGCGTCANGCTATGTGCTG
CCCCGGAATTACTGCAAAAATGGAATATGTGTGNTTCTGATCAAAATCATTTCCGAGGAGAAA
TTGAGGAAACCATCACTGAAAGCTTTGGTAATGATCATAGCACCTTGGATGGG

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCTGGTAGTCGCGCGGTGTGGCTGCACCTCACCAATCCCGTGCGCCGCGGCTG
GGCCGTCGGGAGAGTGCCTGTGCTTCTCTCCTGCACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCGCGCCAGGTTT
GAGGATGGGGGAGTAGCTACAGGAAGCGACCCCGCATGGCAAGGTATATTTTGTGGAATGAAAAGGAAGTATTAGA
AATGAGCTGAAGACCATTACAGATTAATATTTTTGGGGACAGATTTGTGATGCTTGATTACCCCTGAAGTAATGTA
GACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTTCACTTAAATCAGAACTTG
CATAAGAAAGAGAATGGGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAAAGGATCATTCTCTGTTTTCT
GATAGTGTATATGGCCATTTTAGTGGGCACAGATCAGGATTTTACAGTTTACTTGGAGTGTCCAAAACCTGCAAGCAG
TAGAGAAATAAGACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAAAAAACCCGAATAACCCAAATGCACA
TGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGATCTACGGA AAAAGTATGACAAATATGG
AGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAACATATATCGTTATGATTTTGGTATTTATGA
TGATGATCCTGAAATCATAACATTGGAAGAAGAGAATTTGATGCTGCTGTTAATTTCTGGAGAACTGTGGTTTGTAAA
TTTTTACTCCCCAGGCTGTTACACTGCCATGATTAGCTCCACATGGAGAGACTTTGCTAAAGAAGTGGATGGGTT
ACTTCGAATTTGGAGCTGTAACTGTGGTGTATGATAGAATGCTTTGCCGAATGAAAGGAGTCAACAGCTATCCAGTCT
CTTCATTTTTTCGGTCTGGAATGGCCCCAGTGAATATCATGGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTGCAAT
GCAGCATGTTAGAAGTACAGTGACAGAACTTTGGACAGGAAATTTTGTCAACTCCATACAAACTGCTTTTGTCTGCTGG
TATTGGCTGGCTGATCACTTTTTGTTCAAAAGGAGGAGATTGTTTGACTTCACAGACACGACTCAGGCTTAGTGGCAT
GTTGTTTTCTCAACTCATTTGGATGCTAAAGAAATATATTTGGAAAGTAATACATAATCTTCCAGATTTTGAACACTTTT
GGCAACACACTAGAGGATCGTTTGGCTCATCATCGGTGGCTGTTATTTTTTCATTTTGGAAAAAATGAAAATTCAAA
TGATCCTGAGCTGAAAAAACTAAAACTCTACTTAAAAATGATCATATCAAGTTGGCAGGTTTGAAGTGTTCCTCTGC
ACCAGACATCTGTAGTAATCTGTATGTTTTTCAGCCGCTCTCTAGCAGTATTTAAAGGACAAGGAACCAAGAATATGA
AATTCATCATGGAAGAAGATTCTATATGATATACTTGGCTTTGCCAAAGAAAGTGTGAATTTCTATGTTACCAAGCT
TGGACCTCAAAATTTTCTGCCAATGACAAAGAACCATGGCTTGTGATTTCTTTGCCCCCTGGTGTCCACCATGTCTG
AGCTTTACTACAGAGTTACGAAGAGCATCAAATCTTCTTTATGGTCAGCTTAAGTTTGGTACACTAGATTGTACAGT
TCATGAGGGACTCTGTAACATGTATAACATTGAGGCTTATCCAACAACAGTGGTATTTCAACCAGTCCAACATTCATGA
GTATGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ACCCACCACCTTCAACGAACCTAGTTACACAAAGAAAAACACAACGAAGCTGGATGGTTGATTTCTATTCTCCGTGGTG
TCATCCTTGCCAAGTCTTAATGCCAGAATGGAAGAATGGCCCGGACATTAAGTGGACTGATCAACGTGGGCAGTAT
AGATTGCCAACAGTATCATTTCTTTTGTGCCCAGGAAAACGTTCAAAGATACCTTGAGATAAGATTTTTTCCCCAAA
ATCAAATAAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGGATGCTTATTCCTGAGAATCTGGGGTCTAGG
ATTTTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTCAGTGAAAAAGTTCTACAAGGGAAAAATCATTGGGT
GATTGATTTCTATGCTCTGCTTGGTGTGGACCTTGCCGAATTTTGCTCCAGAAATTTGAGCTCTTGGCTAGGATGATTAA
AGGAAAAGTGAAGCTGGAAGAAGTAGACTGTGAGGCTTATGCTCAGACATGCCAGAAAGCTGGGATCAGGGCCTATCC
AACTGTTAAGTTTTATTTCTACGAAAGAGCAAAGAGAAATTTTCAAGAAGACGAGATAAATACCAGAGATGCAAAAGC
AATCGCTGCTTAAATAAGTGA AAAAATTGGA AACTCTCCGAATCAAGGCAAGAGGAATAAGGATGAACTTTGATAATG
TTGAAGATGAAGAAAAAGTTTAAAAGAAATTCTGACAGATGACATCAGAAGACACCTATTTAGAATGTTACATTTATG
ATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGTACTGCCAGAATTATCTACAGCACTGGTGTAAAAGAAGGGT
CTGCAAACTTTTTCTGTAAAGGGCCGGTTTATAAATATTTTAGACTTTGCAGGCTATAATATATGGTTCACACATGAG
AACAAGAATAGAGTCATCATGTATTCTTTGTTATTTGCTTTTAAACAACCTTTAAAAAATATTA AAACGATTCTTAGCT
CAGAGCCATACAAAAGTAGGCTGGATTGAGTCCATGGACCATAGATTGCTGTCCCTCGACGGACTTATAATGTTTC
AGGTGGCTGGCTTGAACATGAGTCTGCTGTGCTATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTTCCCTTAC
GTTTTTTGGCTGACCTGAAAAGAGGTAACTTAGTTTTTGGTCACTTGTCTCTCTAAAAATGCTATCCCTAACCATATA
TTTATATTTCTGTTTTTAAAAACACCCATGATGTGGCACAGTAAACAAACCTGTTATGCTGTATTATTATGAGGAGATT
CTTCATTGTTTTCTTTCCCTTCTCA
AAGTTGAAAAAATGCTTTTAAATTTTTCACAGCCGAGAAACAGTGCAGCAGTATATGTGCACACAGTAAGTACAC
AAATTTGAGCAACAGTAAGTGCACAAATCTGTAGTTTGTGTATCATCCAGGAAAAACCTGAGGGAAAAAATTA
TAGCAATTAACGGGCATTGTAGAGTATCCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATATGT
GTTTCATGTATTTTCTGAAATGCTTTTCATAGAAATTTTCCCACTGATAGTTGATTTTGGAGCATCTAATATTTACAT
ATTTGCCTTCTGAACTTTGTTTGACCTGTATCCTTTATTTACATTGGGTTTTTCTTTTCATAGTTTTGGTTTTTCACT
CCTGTCCAGTCTATTTATTTATTTCAAATAGGAAAAATTACTTTACAGGTTGTTTTACTGTAGCTTATAATGATACTGTA
GTTATTCCAGTTACTAGTTTACTGTGAGAGGGCTGCCTTTTTTTCAGATAAATATTGACATAATAACTGAAGTTATTTTT
ATAAGAAAATCAAGTATATAAATCTAGGAAAGGGATCTTCTAGTTTCTGTGTTGTTTAGACTCAAAGAATCAAAATT
TGTCAGTAACATGTAGTTGTTTAGTTATAATTGAGAGTGACAGAAATGGTAAAAATCCAATCAGTCAAAAGAGGTCA
ATGAATTAAGGCTGCACTTTTTTCAAAAAA AAAAAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKLHPD
KNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLLEDNQGGQYESWNYRYDFGIYDDDE
IITLERREFDAAVNSGELWVFNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNCGDDRMLCRM
KGVNSYPSLFI FRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNSIQTAFAAGIGWL
ITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSANTLEDRLAHRWLLFF
HFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSNLYVFQPSLAVFKGQGTKEYEIIHG
KKILYDILAFAKESVNSHVTTLGPQNFPANDKEPWLVDFFAPWCPPCRALLPELRRASNLLYGQL
KFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEHEGHHSAEQILEFIEDLMNPSVVS LTPPTF
NELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMARTLTGLINVGSIDCQQYHSFCAQENVQRY
PEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLGFLPQVSTD LTPQTFSEKVLQGNHWVIDFY
APWCGPCQNFAPFELLARMIKGKVKAGKVDCQAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEE
QINTRDAKAIAALISEKLETLRNQGKRKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAAGCC
ATGAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGAGTCGTT
GGTGAAGTTTTTCATTCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCATTACTGGAG
CTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGCATATTGGTTCTG
TGGGATATTAATAAGCGCGGTGTGGAGGAACTGCAGCTGAGTGCCGAAAAC TAGGCGTCACTGC
GCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCTCTCTAAATCAGGTGAAGA
AAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACAGTATATCCAGCCGATCTTCTC
AGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAACATCCTAGGACATTTTTGGATCAC
AAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATGGCCACATCGTCACAGTGGCTTCAGTGT
GCGGCCACGAAGGGATTCTTACCTCATCCCATATTGTTCCAGCAAATTTGCCGCTGTTGGCTTT
CACAGAGGTCTGACATCAGAACTTCAGGCCTTGGGAAAAACTGGTATCAAAACCTCATGTCTCTG
CCCAGTTTTTGTGAATACTGGGTTACCAAAAATCCAAGCACAAGATTATGGCCTGTATTGGAGA
CAGATGAAGTCGTAAGAAGTCTGATAGATGGAATACTTACCAATAAGAAAATGATTTTTGTTCCA
TCGTATATCAATATCTTTCTGAGACTACAGAAGTTTCTTCCTGAACGCGCCTCAGCGATTTTAAA
TCGTATGCAGAATATTCAATTTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAATGAATAAATA
AGCTCCAGCCAGAGATGTATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAAGCTT
TATTTACATTTTTTTCAGTCCTGATAATATTA AAAACATTGGTTTGGCACTAGCAGCAGTCAAAC
GAACAAGATTAATTACCTGTCTTCCTGTTTCTCAAGAATATTTACGTAGTTTTTTCATAGGTCTGT
TTTTCTTTTCATGCCTCTTAAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTCTTTAA
GATATTTTATTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAA
CTTATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTTCCAATCTGTAGCCATGCCACAGA
ATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTTATCT
CAACCTGGACATATTTTAAGATTCAGCATTTGAAAGATTTCCCTAGCCTCTTCCTTTTTTCATTAG
CCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTATAACTCTGAAG
TCCACCAAAGTGGACCCTCTATATTTCCCTCCCTTTTTATAGTCTTATAAGATACATTATGAAAG
GTGACCGACTCTATTTTAAATCTCAGAAATTTAAGTTCTAGCCCCATGATAACCTTTTTCTTTGT
AATTTATGCTTTCATATATCCTTGGTCCAGAGATGTTTAGACAATTTTAGGCTCAAAAATTAAA
GCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAACAATGGACCCAAGAGAAGAA

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTFYFAKRQSILVL
WDINKRGVEETAAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVVNNAGTVYPADLL
STKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYCSSKFAAVGF
HRGLTSELQALGKTGIKTSCLCPVFNVTGFTKNPSTRLWPVLETDEVVRSLIDGILTNNKMI FVP
SYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGGATG
CCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCCAGCCCGCCCGGGGCAGGATG
ACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTTCATGATCCTGCTGATCAT
CGTGTA TGGGACAGCGCAGGCGCCGCGCACTTCTACTTGACACGTCCTTCTCTAGGCCGCACA
CGGGGCCGCGCTGCCCACGCCCAGGGCCGGACAGGGACAGGGAGCTCACGGCCGACTCCGATGTC
GACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGACCTTCCCAGAAAGGAGAC
GGAGCAGCCGCCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAGGCTACGACTGGTCCCCGCGCG
ACGCCCCGGCGCAGCCCAGACCAGGGCCGGCAGCAGGCGGAGCGGAGGAGCGTGCTGCGGGGCTTC
TGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCGCGCATTTCGACGACATCCCCAACTCGGA
GCTGAGCCACCTGATCGTGACGACCGGCACGGGGCCATCTACTGCTACGTGCCCAAGGTGGCCT
GCACCAACTGGAAGCGCGTGATGATCGTGCTGAGCGGAAGCCTGCTGCACCGCGGTGCGCCCTAC
CGCGACCCGCTGCGCATCCCGCGCGAGCACGTGCACAACGCCAGCGCGCACCTGACCTTCAACAA
GTTCTGGCGCCGCTACGGGAAGCTCTCCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCA
AGTTCCTCTTCGTGCGCGACCCCTTCGTGCGCCTGATCTCCGCCTTCCGCAGCAAGTTCGAGCTG
GAGAACGAGGAGTTCTACCGCAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACCACACCAG
CCTGCCCCGCTCGGCGCGCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCCTTCGCCAACTTCATCC
AGTACCTGCTGGACCCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTAC
CGCCTCTGCCACCCGTGCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGA
CGCCGCGCAGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGA
ACAGGACCGCCAGCAGCTGGGAGGAGGACTGGTTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAG
CTGTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCCG
AGACTGAAGCTTTCGCGTTGCTTTTTTCTCGCGTGCTGGAACCTGACGCACGCGCACTCCAGTT
TTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCCATTGAGT
ACTGTATCGATATTGTTTTTTAAGATTAATATATTTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSFSPHTGPPLPTPGPDRDRELTADSD
VDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAERRSVLRG
FCANSSLAFTKERAFFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRMIVLSGSLLRGAP
YRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPFVRLISAFRSKFE
LENEEFYRKFAVPMRLRYANHTSLPASAREAFRAGLKVSFANFIQYLLDPHTEKLAPFNEHWRQV
YRLCHPCQIDYDFVGKLETLDDEAAQLLQLLQVDRQLRFPPSYRNRTASSWEEDWFAKIPLAWRQ
QLYKLYEADFVLFGYPKPENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

TCGGGCCAGAATTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGAAAG
AGGCCCAGAGTAGAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCTATGCCG
GGAAGGTGGTGGTCGTGACCGGGGGCGGGCGGCATCGGAGCTGGGATCGTGCGCGCCTTCGTG
AACAGCGGGGCCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCGGGCCCTGGAGCAGGA
GCTCCCTGGAGCTGTCCTTATCCTCTGTGATGTGACTCAGGAAGATGATGTGAAGACCCTGGTTT
CTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAACGCTGGCCACCACCCACCC
CCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCTGCTGGAGCTGAACCTACTGGG
GACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGAAGAGTCAAGGGAATGTCATCAACA
TCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCAGTTCCTATGTGGCCACCAAGGGGGCA
GTAACAGCCATGACCAAAGCTTTGGCCCTGGATGAAAGTCCATATGGTGTCCGAGTCAACTGTAT
CTCCCCAGGAAACATCTGGACCCCGCTGTGGGAGGAGCTGGCAGCCTTAATGCCAGACCCTAGGG
CCACAATCCGAGAGGGCATGCTGGCCCAGCCACTGGGCCGCATGGGCCAGCCCGCTGAGGTCGGG
GCTGCGGCAGTGTTCCCTGGCCTCCGAAGCCAACCTTCTGCACGGGCATTGAACTGCTCGTGACGGG
GGGTGCAGAGCTGGGGTACGGGTGCAAGGCCAGTCGGAGCACCCCCGTGGACGCCCCCGATATCC
CTTCCTGATTTCTCTCATTTCTACTTGGGGCCCCCTTCCTAGGACTCTCCCACCCCAAACCTCAA
CCTGTATCAGATGCAGCCCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTCACC
CTGCAGGTTCCCATAAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045

<subunit 1 of 1, 270 aa, 1 stop

<MW: 28317, pI: 6.00, NX(S/T): 1

MATGTRYAGKVVVVTGGGRGIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCDVTQ
EDDVKTLVSETIRRFGRLCDVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLTKLALPYLR
KSQGNVINISLVAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGNIWTPLWEEL
AALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTGGAELGYGCKASRS
TPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAACTGCTG
TTTCTTCTTACCATTTCCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAAGAGGAA
GGGGCAAGGGCGGCCTGGGCCCCTGGCCCCTGGCCCCTCACCAGGTGCCACTGGACCTGGTGTAC
GGATGAAACCGTATGCCCCGATGGAGGAGTATGAGAGGAACATCGAGGAGATGGTGGCCCAGCTG
AGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCTGTGGATGTCCAACAA
GAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCCGTATCCCCGTGGACCTGC
CGGAGGCACGGTGCTGTGTCTGGGCTGTGTGAACCCCTTCACCATGCAGGAGGACCGCAGCATG
GTGAGCGTGCCGGTGTTCAGCCAGGTTCTGTGCGCCGCCGCTCTGCCCCGCCACCGCCCCGCAC
AGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCGCTGTGGGCTGCACCTGCATCTTCTTGAA
TCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGACCATCCTCCTTGACCTTTGTGCCAAGAAA
GGCCTATGAAAAGTAAACACTGACTTTTGAAAGCAAG

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEYERN

IEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSPRIPVDLPEARCLCLGCVNPF

TMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGGCGA
GCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCGGCGCCC
AACATGGCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCTGGATCGCGGCTGT
GGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGAGCAGAGCCGGGTCCAGCCCATGA
CCGCCTCCAACCTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTTTACGCCCCATGGTGT
CCATCCTGCCAGCAGACTGATTCAAGATGGGAGGCTTTTGCAAAGAATGGTGAAATACTTCAGAT
CAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTGGCCGCTTCTTTGTCACCACTC
TCCCAGCATTTTTTTCATGCAAAGGATGGGATATTCCGCCGTTATCGTGGCCCAGGAATCTTCGAA
GACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATCAGTCGAGCCTCTGACTGGCTGGAAATC
CCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTCTTTTTCAGCATCTCTGGCAAGATATGGCATC
TTCACAACATATTTTCACAGTGACTCTTGAATTCCTGCTTGGTGTTCTTATGTGTTTTTCGTCATA
GCCACCTTGGTTTTTGGCCTTTTTATGGGTCTGGTCTTGGTGGTAATATCAGAATGTTTCTATGT
GCCACTTCCAAGGCATTTATCTGAGCGTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAG
CTGAACAGTTGCAGGATGCGGAGGAGGAAAAAGATGATTCAAATGAAGAAGAAAAACAAAGACAGC
CTTGTAGATGATGAAGAAGAGAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGA
GGAGGACAACCTTGGCTGCTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCAG
GAGAGGACGGTGTGACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAA
CCCTGCCCAGCTGACACAGAGGTGGTGAAGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGA
CAAGGGACTGTAGATTTAATGATGCGTTTTCAAGAATACACACCAAACAATATGTCAGCTTCCC
TTTGGCCTGCAGTTTGTACCAAATCCTTAATTTTTCTGAATGAGCAAGCTTCTCTTAAAAGATG
CTCTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGTG
ACAATCAGGATATAGAAAAACAAACGTAAGTGTGGGATCTGTTTGGAGACTGGGATGGGAACAAG
TTCATTTACTTAGGGGTGAGAGAGTCTCGACCAGAGGAGGCCATTCCAGTCCTAATCAGCACCT
TCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCTCCTGAGCATC
CCCAAAGTGTAACGTAGAAGCCTTGCATCCTTTTCTGTGTAAAGTATTTATTTTTGTCAAATTG
CAGGAAACATCAGGCACCACAGTGTCATGAAAAATCTTTCACAGCTAGAAATTGAAAGGGCCTTGG
GTATAGAGAGCAGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTGTGCTATGTTTTATTTCTTA
CCTTTAATTTTTTCCAGCATTTCCACCATGGGCATTGAGGCTCTCCACACTCTTCACTATTATCTC
TTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAACTGTGTTTGTTCATTCTGACCTAAGG
GGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGTGACTGCCAAACATCTCAAATGAAATGTT
GTGGCCATCAGAGACTCAAAGGAAGTAAGGATTTTACAAGACAGATTAAAAAAAATTTGTTTTG
TCCAAAATATAGTTGTTGTTGATTTTTTTTTTAAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGT
CCTCTAAGTCTTGCCAGTACAAGGTAGTCTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTCATC
TCAAGGGGTTCCCTGGGTCTTGAACACTTTTAATAATAACTAAAAAACCACTTCTGATTTTCCTT
CAGTGATGTGCTTTTGGTGAAAGAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTTGATTTT
GTTTCCATCTTCTGTAATCTTCCAAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACT
GTAAGTACCCAGGAGGCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAPWCP
SCQQTDSWEAFKNGEILQISVGKVDVIEPGLSGRFFVTTLPAFFHAKDGIFRRYRGP GIGFED
LQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAWCSYVFFVIA
TLVFGLFMGLVLVVISSECFYVPLPRHLSESEQNRRSEEAHRAEQLQDAEEEEKDDSNEEENKDSL
VDDEEEKEDLGDEDEAEEDNLAAGVDEERSEANDQPPGEDGVTREEVEPEEAEEGISEQP
CPADTEVVEDSLRQRKSQHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

FIGURE 201

ATCTGGTTGAACTACTTAAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATTTGA
CTCAGAGATTCTCTTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCCCAAATG
CTATATCTATTTCAGGGGCTCTCAAGAACTAATGGAATATCATCCTGATTTAGAAAATTTGGATGAA
GATGGATATACTCAATTACACTTTCGACTCTCAAAGCAATACCAGGATAGCTGTTGTTTCAGAGAA
AGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCTGTAATTTTGGGAATCCTATGCTTGG
TAATACTGGTGATAGCTGTGGTCCCTGGGTACCATGGGGGTTCTTTCCAGCCCTTGTCTCCTAAT
TGGATTATATATGAGAAGAGCTGTTATCTATTTCAGCATGTCACTAAATTCCTGGGATGGAAGTAA
AAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAGACAGCTCAAATGAATTGGGATTTA
TAGTAAAACAAGTGTCTTCCCAACCTGATAATTCATTTTGGATAGGCCTTCTCGGCCCCAGACT
GAGGTACCATGGCTCTGGGAGGATGGATCAACATTCTCTTCTAACTTATTTTCAGATCAGAACAC
AGCTACCCAAGAAAACCCATCTCCAAATTGTGTATGGATTACGTGTCAGTCATTTATGACCAAC
TGTGTAGTGTGCCCTCATATAGTATTTGTGAGAAGAAGTTTTCAATGTAAAGAGGAAGGGTGGAGA
AGGAGAGAGAAAATATGTGAGGTAGTAAGGAGGACAGAAAACAGAACAGAAAGATTAACAGCTGA
GGTCAAGATAAATGCAGAAAATGTTTAGAGAGCTTGGCCAACGTGAATCTTAACCAAGAAATTGA
AGGGAGAGGCTGTGATTTCTGTATTTGTGCGACCTACAGGTAGGCTAGTATTATTTTTCTAGTTAG
TAGATCCCTAGACATGGAATCAGGGCAGCCAAGCTTGAGTTTTTTATTTTTTATTTATTTATTTTT
TTGAGATAGGGTCTCACTTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAG
CTATCTCTCGCCTCAGCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAAT
TTTTGGTGTTTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGC
TTAAGTGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTG
GCCCCAAGCTTGAATTTTCTATTCTGCCATTGACTTGGCATTTACCTTGGGTAAAGCCATAAGCGAA
TCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCACGGTGTGT
TGCCACGATTTGACCTCAACTTCTAGCAGTATATCAGTTATGAACTGAGGGTGAAATATATTTTC
TGAATAGCTAAATGAAGAAATGGGAAAAAATCTTCACCACAGTCAGAGCAATTTTATTATTTTCA
TCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAACCTCTACTTTTTCTTTATC
AATTAAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTTTTTTTTTTTTTTTTTTT
TTTGAGACAGAGTTTCGCTCTTGTTGCCAGGCTGGAGTGCAACGGCACGATCTCGGCTCACCGC
AACCTCCGCCCCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGT
CAGGCACCACCACACCCGGCTAATTTTGTATTTTTTTTAGTAGAGACAGGGTTTCTCCATGTCCGT
CAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGCCTGCCTCGGCCTCCCAAGTGCTGGGATT
ACAGGCGTGAGCCACTGCACCCAGCCTAGAATCTTGATAATATGTAATTGTAGGGAACTGCTC
TCATAGGAAAGTTTTCTGCTTTTTAAATACAAAAATACATAAAAAATACATAAAATCTGATGATGA
ATATAAAAAAGTAACCAACCTCATTGGAACAAGTATTAACATTTTGAATATGTTTTATTAGTTT
TGTGATGTACTGTTTTACAATTTTACCATTTTTTTTCAGTAATTACTGTAAAATGGTATTATTGG
AATGAAACTATATTTCTCATGTGCTGATTTGTCTTATTTTTTTTCATACTTTCCCACTGGTGCTA
TTTTTATTTCCAATGGATATTTCTGTATTACTAGGGAGGCATTTACAGTCCTCTAATGTTGATTA
ATATGTGAAAAGAAATTGTACCAATTTTACTAAATTATGCAGTTTAAAATGGATGATTTTATGTT
ATGTGGATTTCAATTTCAATAAAAAAAACTCTTATCAAAAAAATAAAAAAAAAAAAAAAAAAAAA
AAAAAAA

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSAAASPPWRLIAVILGILCLVILVIAVVLG
TMGVLSSPCPPNWIIYEKSCYLFMSLSNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQVSSQPD
NSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQLCSVPSYSIC
EKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

FIGURE 203

[illegible]

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921

<subunit 1 of 1, 693 aa, 1 stop

<MW: 77738, pI: 8.87, NX(S/T): 7

MTPQSLQLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIENSEEAL
TVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLSDKASSLLCFQHQEESLA
QGPPLLATSVTSWWSPQNI SLPSAASFTFSFHSPPHTAAHNASVDMCELKRDQLLSQFLKHPQK
ASRRPSAAPASQQQLQSLESKLTSVRFMGDMVVSFEEDRINATVWKLQPTAGLQDLHIHSRQEEEQS
EIMEYSVLLPRTLFRQTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGEKVLGIVVQNTKVANL
TEPVVLTFQHQLQPKNVTLQCVFWVEDPTLSSPGHWSSAGCETVRRETQTSCFCNHLTYFAVLMV
SSVEVDAVHKHYLSLLSYVGCVVSAALACLVITIAAYLCSRVP LPCR KPRDYTIKVHMNLLLA VFL
LDTSFLLSEPVALTGSEAGCRASAI FLHFSLLTCLSWMGLEGYNLYRLVVEVFGTYVPGYLLKLS
AMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIYPSMCWIRDSLVSYITNLGLFSLVFLFN
MAMLATMVVQILRLRPHTQKWSHVLTLGLSLVLGLPWALIFFSFASGTFQLVVLYLEFSIITSFQ
GFLIFIWYWSMRLQARGGPSPLKSNSDSARLPISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

FIGURE 205

TGCCTGGCCTGCCTTGTCAACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGAGGA
AANCNTCGGGACTACACCNTCAAGTGCACATGAACCTGCTGCTGGCCGTCTTCCTGCTGGACACG
AGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGAGCCAGTGCC
ATCTTCCTGCACTTCTCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGGTACAACCTCTAC
CGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGG
CTGGGGCTTCCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACCTATGGCCCCA
TCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCTTCCATGTGCTGGATCCGGGAC
TCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCTGGTGTTCCTGTTCAACATGG

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTCAGGTCCAGGTTTTGCTTTGATCC
TTTTCAAAAAGCTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAACTACCTGCGGATT
CTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCCAAGTGCAGCCTTCCCCTGGCGGTGGTGAAAGAGACTCGGGAGTC
GCTGCTTCCAAAGTGCCCCGCCGTGAGTGAGCTCTACCCCAAGTGCAGCCTTCCGCGCTTCTCCTGCTGA
CATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCAGTTTTCCAGCAACA
AGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATGGAAGTATTACAGCCCAA
GGTTTCCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAGAGGAAATGTATGGATACAAC
TTACGTTTGTATGAAAGATTGGGCTTGAAGACCCAGAAGATGACATATGCAAGTATGATTTTGTAGAAGTTGAGGAAC
CCAGTGATGGAATATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAGGAAAAACAGATTTCTAAAGGAAATCAAA
TTAGGATAAGATTTGTATCTGATGAATATTTTCTTCTGAACCAGGGTTCTGCATCCACTACAACATTGTTCATGCCAC
AATTCACAGAAGCTGTGAGTCCTTCAGTGCTACCCCTTCAGCTTTGCCACTGGACCTGCTTAATAATGCTATAACTG
CCTTTAGTACCTTGGAAAGACCTTATTCGATATCTTGAACCAGAGAGATGGCAGTTGGACTTAGAAGATCTATATAGGC
CAACTTGGCAACTTCTTGGCAAGGCTTTTGTTTTTTGAAGAAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGG
AGGTAAGATTATACAGCTGCACACCTCGTAACCTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTT
TCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAAT
GTGTCCCAAGCAAGTTACTAAAAAATACCACGAGGTCTTCAGTTGAGACCAAAGACCGGTGTGAGGGGATTGCACA
AATCACTCACCGACGTGGCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGTCAGAGGGAGCACAGGAGGATAGCCGC
ATCACCACCAGCAGCTCTTGCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCAT
CCTTAATCTCAGTTGTTTGTCTCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAAC
AGAATTAGGAGTTGTGCAACAGCTCTTTTGAAGAGGAGGCCTAAAGGACAGGAGAAAAAGGCTTCAATCGTGGAAAGAA
AATTAAATGTTGTATTAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCTAGCTGGGTTCTGTAT
TTCAGTTCTTTTCGATACGGCTTAGGGTAATGTGAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCC
TTGCTTAACTCTAAAGCTCCATGTCTTGGCCTAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTTTGTCTCATAT
TCACATATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAACCT
TGTGTATGCTGATAGGACAGACTGGATTTTTTCATATTTCTTATTAATTTCTGCCATTTAGAAGAAGAGAACTACA
TTCATGGTTTTGGAAGAGATAAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTGTTT
CATTGTGTACATTTTATATTTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTAAATATATCTATTTTT
ACCAAAGGTATTTAATATCTTTTTTATGACAACCTTAGATCAACTATTTTAGCTTGGTAAATTTTTCTAAACACAAT
TGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCAATCTCGTATGGT
GCTAGAGTTAGATTAATCTGCATTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAAAATA
ATTAATTTATCATATCTTCCATTCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTGATCC
AGCCATTACTAACCTATTCCTTTTTTGGGGAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTT
GGCAGCTTCCTGATAAAGCGTGTCTGTGCTGTGTCAGTAGGAACACATCCTATTTATTGTGATGTTGTGGTTTTATTATC
TTAAACTCTGTTCCATACACTTGTATAAATACATGGATATTTTTATGTACAGAAGTATGTCTCTTAACCAGTTCACTT
ATTGTACTCTGGCAATTTAAAAAGAAAATCAGTAAAAATTTTTGCTTGTAAAAATGCTTAATATNGTGCCTAGGTTATGT
GGTGAATTTGAATCAAAAATGTATTGAATCATCAATAAAAGAAATGTGGCTATTTTGGGGAGAAAATTAATAAAAAA
AAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

FIGURE 207

MSLFGLLLLTALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIIITVSTNGSIHSPRFPH
TYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWCGSGTVP
GKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLLNNAITAFST
LEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLEEVRLYSCTPRNFSVSIR
EELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQLRPKTGVRGLHKSLT
DVALEHHEECDCVCRGSTGG

Signal sequence:

amino acids 1-14

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGGAAGAAGA
CTAAAAATGGTGTGTTTCCAATGTGGACACTGAAGAGACAAATCTTATCCTTTTAAACATAATCCTAATTTCCAAACTC
CTTGGGGCTAGATGGTTTCTTCTAAAACTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTGATCGTGGAC
TGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTTCCACGAACACCACGAACCTCACCTCACCATTAAACCAC
ATACCAGACATCTCCCCAGCGTCCTTTACAGACTGGACCATCTGGTAGAGATCGATTTTACAGATGCAACTGTGTACCT
ATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGCTTTAGTGGACTCACTTAT
TTAAAAATCCCTTTACCTGGATGGAAACCAGCTACTAGAGATACCGCAGGGCCTCCCGCCTAGCTTACAGCTTCTCAGC
CTTGAGGCCAACACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCCAAATAGAAATACTCTACCTGGGC
CAAACTGTTATTATCGAAATCCTTGTTATGTTTCATATTCAATAGAGAAAGATGCCTTCTTAAACTTGACAAAGTTA
AAAGTGCTCTCCCTGAAAGATAACAATGTACAGCCGTCCTACTGTTTGGCCATCTACTTTAACAGAACTATATCTC
TACAACAACATGATGTGCAAAAATCCAAGAAGATGATTTTAATAACCTCAACCAATTACAAATTCTTGACCTAAGTGA
AATTGCCCTCGTTGTTATAATGCCCCATTTCTTGTGCGCCGTGTAATAATAATTCTCCCTACAGATCCCTGTAAAT
GCTTTTGATGCGCTGACAGAATTAAGGTTTTACGTCTACACAGTAACTCTCTTCAGCATGTGCCCCAAGATGGTTT
AAGAACATCAACAACTCCAGGAAGTGGATCTGTCCCAAACTTCTTGGCCAAAGAAATTGGGGATGCTAAATTTCTG
CATTTTCTCCCCAGCCTCATCCAATTGGATCTGTCTTTCAATTTGAACTTCAGGTCTATCGTGCATCTATGAATCTA
TCACAAGCATTTTCTTCACTGAAAAGCCTGAAAATTTCTGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTT
AACCTCTCGCCATTACATAATCTTCAAAATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAAAATTGCTAACCTC
AGCATGTTTAAACAATTTAAAAGACTGAAAGTCATAGATCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGT
GAAGTTGGCTTCTGCTCAAATGCCAGAACTTCTGTAGAAAGTTATGAACCCAGGTCTGGAACAATTACATTATTTT
AGATATGATAAGTATGCAAGGAGTTGCAGATTCAAAAACAAGAGGCTTCTTTCATGTCTGTTAATGAAAGCTGCTAC
AAGTATGGGCAGACCTTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCCTCTGATTTTCAGCATCTTTCTTT
CTCAAATGCCCTGAATCTGTGAGGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATTCCAACCTTTAGCAGAGCTG
AGATATTTGGACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAGAGCTTCACAACTGGAAGTT
CTGGATATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAACTTTACCAAGAACCTAAAG
GTTCTGCAGAACTGATGATGAACGACAATGACATCTCTTCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
ACTCTGGAATTCAGAGGAAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAGAAT
CTGCTAAAATTAGAGGAATTAGACATCTCTAAAAATTCCTAAGTTTCTTGCTTCTGAGTTTTTGATGGTATGCCT
CCAAATCTAAAGAATCTCTCTTTGGCCAAAAATGGGCTCAAATCTTTCAGTTGGAAGAACTCCAGTGTCTAAAGAAC
CTGGAACCTTTGGACCTCAGCCACAACCACTGACCACTGTCCCTGAGAGATTATCCAAGTGTTCAGAAGCCTCAAG
AATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTCCAGTTGCGATATCTG
GATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCCAGAAAATGTCTCAACAATCTGAAGATGTTG
CTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTGTCTGGTGGGTTAACCATACGGAGGTGACT
ATTCCTTACCTGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGCCAAAGTGTGATCTCCCTGGATCTG
TACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATATCTGTATCTCTCTTCTCATGGTGATG
ATGACAGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCATTCTGTAAAGGCCAAGATAAAGGGGTATCAG
CGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTATGACACTAAAGACCCAGCTGTGACCGAGTGGGTT
TTGGCTGAGCTGGTGGCCAACTGGAAGACCCAAGAGAGAAACATTTTAATTTATGTCTCGAGGAAAGGGACTGGTTA
CCAGGGCAGCCAGTTCTGGAACCTTTCCCAGAGCATACAGCTTAGCAAAAAGACAGTGTGTTGTGATGACAGACAA
TATGCAAGACTGAAAATTTTAAGATAGCATTTTACTTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATT
ATCTTGATATTTCTTGAGAAGCCCTTTTCAAGTCCAAGTTCCTCCAGCTCCGAAAAGGCTCTGTGGGAGTTCTGTC
CTTGAGTGGCCAAACCCGCAAGCTCACCCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCAT
GTGGCCTATAGTCAGGTGTTCAAGGAAACGGTCTAGCCCTTCTTTGCAAAACACAACCTGCCTAGTTTACCAAGGAGAG
GCCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGGIPT
NTTNLTTLTINHIPDISPASFHRLDHLVEIDFRNCNCPVPIPLGSKNNMCIKRLQIKPRSFSGLTYLK
SLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYYRNPCYVSYSI
EKDAFLNLTCLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNLNQLQILDLSGNC
PRCYNAPFFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWFKNINKLQELDLSQNF
LAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSSLKSLKILRIRGYVFKELKSFNL
SPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKISPSGDSSEVGFCSNARTSVESY
EPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQTLDSLKNISIFFVKSSDFQHLSFLK
CLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRDLLHSTAFEELHKLEVLDISSNSHYFQSEGI
THMLNFTKNLKVQLKMMNDNDISSSTSRTMESESLRTLEFRGNHLDVLWREGDNRYLQLFKNLL
KLEELDISKNLSFLPSGVFDGMPPNLKNLSLAKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVP
ERLSNCSRSLKNLILKNNQIRSLTKYFLQDAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLL
HHNRFLCTCDAVWVFWVWNHTEVTIPYLATDVTCTVGPGAHKGQSVISLDLYTCELDLTNLILFSL
SISVSLFLMVMMTASHLYFWDVWYIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPAVTEWVLA
ELVAKLEDPREKHFNLCLERDWLPGQPVLNLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSH
QRLMDEKVDVILIFLEKPFQKSKFLQLRKRLCGSSVLEWPTNPQAHFYFWQCLKNALATDNHVA
YSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAACAGAAACATGGAAAAATGTTCCCTTCAGT
CGTCAATGCTGACCTGCATTTTCTGCTAATATCTGGTTCCCTGTGAGTTATGCGCCGAAGAAAAATTTTCTAGAAAGCT
ATCCTTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGCAGAGTGCAGCAATCGTCTGACTACAGGAAGTTCCCCAAA
CGGTGGGCAAAATATGTGACAGAACTAGACCTGTCTGATAATTTATCACACACATAACGAATGAATCATTTCAGGGC
TGCAAAATCTCACTAAAATAAATCTAAACCACAACCCCAATGTACAGCACCAGAACGGAAATCCCGGTATACAATCAA
ATGGCTTGAATATCACAGACGGGGCATTCTCAACCTAAAAAACCTAAGGGAGTTACTGCTTGAAGACAACCAGTTAC
CCAAAATACCCCTCTGGTTTGGCCAGAGTCTTTGACAGAACTTAGTCTAATTCAAAAACAATATATACAACATAACTAAAG
AGGGCATTTCAGACTTATAAACTTGAAAAATCTCTATTTGGCCTGGAAGTCTATTTTAAACAAAGTTTGGCAGAAAA
CTAACATAGAAGATGGAGTATTTGAAACGCTGACAAATTTGGAGTTGCTATCACTATCTTTCAATTCTCTTCACACG
TGCCACCCAACTGCCAAGCTCCCTACGCAAACTTTTTCTGAGCAACACCCAGATCAAATACATTAGTGAAGAAGATT
TCAAGGGATTGATAAATTTAACATTACTAGATTTAAGCGGGAAGTGTCCGAGGTGCTTCAATGCCCATTTCCATGCG
TGCCTTGTGATGGTGGTTCATTAATATAGATCGTTTGTCTTTCAAAAACCTTGACCCAACTTCGATACCTAAACC
TCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCGTGGTTTAAAAATATGCCTCATCTGAAGGTGCTGGATCTTGAAT
TCAACTATTTAGTGGGAGAAATAGTCTCTGGGGCATTTTTAAACGATGCTGCCCGCTTAGAAATACTTGACTTGTCTT
TTAACTATATAAAGGGGAGTTATCCACAGCATATTAATATTTCCAGAACTTCTCTAACTTTTGTCTCTACGGGCAT
TGCAATTTAAGAGGTATGTGTTCCAGGAACCTCAGAGAAGATGATTTCCAGCCCTGATGCAGCTTCCAACTTATCGA
CTATCAACTTGGGTATTAATTTTATTAAGCAAATCGATTTCAAACCTTTTCCAAAATTTCTCCAATCTGGAAATTAATTT
ACTTGTGCAAAAACAGAATATCACCGTTGGTAAAGATACCCGGCAGAGTTATGCAAAATAGTTTCTCTTTTCAACGTC
ATATCCGGAACGACGCTCAACAGATTTTGTAGTTTGAACCCACATTGCAACTTTTATCATTTTACCCGCTCTTTAATAA
AGCCACAATGTGCTGCTTATGGAAGAGCCTTAGATTTAAGCCTCAACAGTATTTTCTTATTGGGCCAAACCAATTTG
AAAATCTTCTGACATTGCCTGTTTAAATCTGTCTGCAAAATAGCAATGCTCAAGTGTTAAGTGGAACTGAATTTTCAG
CCATTCTCATGTCAAATATTTGGATTGACAAACAATAGACTAGACTTTGATAATGCTAGTGCTCTTACTGAATTGT
CCGACTTGGAGTTCTAGATCTCAGCTATAATTCACACTATTTTCAAGATAGCAGGCGTAACACATCATCTAGAATTTA
TTCAAAATTTCAAAAATGATAAAGTTTTAACTTGAGCCACAACAACATTTATACTTTAACAGATAAGTATAACCTGG
AAAGCAAGTCCCTGGTAGAATTTAGTTTTTCAAGTGGCAATCGCCTTGACATTTTGTGGAATGATGACAAACAGGTATA
TCTCCATTTTCAAGGTCTCAAGAATCTGACAGCTCTGGATTATCCCTTAATAGGCTGAAGCACATCCCAAATGAAG
CATTCCTTAATTTGCCAGCGAGTCTCACTGAACATACATATAAATGATAATATGTTAAAGTTTTTTAACTGGACATTAC
TCCAGCAGTTTCTCGTCTCGAGTTGCTTGACTTACGTGGAACAACTACTCTTTTAACTGATAGCCTATCTGACT
TTACATCTTCCCTTCGGACACTGCTGCTGAGTCATAACAGGATTTCCACCTACCCCTCTGGCTTTCTTTCTGAAGTCA
GTAGTCTGAAGCACCTCGATTTAAGTTCCAATCTGCTAAAAACAATCAACAAATCCGCACTTGAACTAAGACCACCA
CCAAATTTATCTATGTTGGAACACACGGAACCCCTTTGAATGCACCTGTGACATTGGAGATTTCCGAAGATGGATGG
ATGAACATCTGAATGTCAAAATTTCCAGACTGGTAGATGTCAATTTGTGCCAGTCTTGGGGATCAAAGAGGGGAAGAGTA
TTGTGAGTCTGAGCTAACAACCTTGTGTTTTCAGATGTCAGTGCAGTGATATTATTTTCTTACGTTCTTTTATCACC
CCATGGTTATGTTGGCTGCCCTGGCTCACCATTGTGTTTACTGGGATGTTTGGTTTATATATATAATGTTTGTAGT
AGGTAAAAGGCTACAGGTCTCTTTCCACATCCCAAACCTTTCTATGATGCTTACATTTCTTATGACACCAAAGATGCCT
CTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACCTGAAGAGAGCCGAGACAAAACGTTCTCCTTTGTCTAG
AGGAGAGGGATTGGGACCCGGGATTGGCCATCATCGACAACCTCATGCAGAGCATCAACCAAAGCAAGAAAACAGTAT
TTGTTTTAACCAAAAAATATGCAAAAAGCTGGAACCTTAAACAGCTTTTACTTGGCTTTGCAGAGGCTAATGGATG
AGAACATGGATGTGATTATATTTATCCTGCTGGAGCCAGTGTACAGCATTCTCAGTATTTGAGGCTACGGCAGCGGA
TCTGTAAGAGCTCCATCCTCCAGTGGCCTGACAACCCGAAGGCAGAAGGCTTGTGTTTGGCAAACCTCTGAGAAATGTGG
TCTTGACTGAAAAATGATTACGGTATAACAATATGTATGTGATTCATTAAAGCAATACTAACTGACGTTAAGTCATG
ATTTTCGCGCCATAATAAAGATGCAAAGGAATGACATTTCTGTATTAGTTATCTATTGCTATGTAACAAATTTATCCCAA
AACTTAGTGGTTTAAACAAACACATTTGCTGGCCACAGTTTTTGGGGTCCAGGAGTCCAGGCCCAGCATACTGGGT
CCTCTGCTCAGGGTGTCTCAGAGGCTGCAATGTAGGTGTTACCCAGAGACATAGGCATCACTGGGGTCACACTCATGT
GGTTGTTTTCTGGATTCAATTCCTCCTGGGCTATTGGCCAAAGGCTATACTCATGTAAGCCATGCGAGCCTCTCCAC
AAGGCAGCTTGTTCATCAGAGCTAGCAAAAAAGAGAGGTTGTAGCAAGATGAAGTCACAATCTTTTGTAAATCGAAT
CAAAAAAGTGATATCTCATCACTTTGGCCATATTCTATTTGTTAGAAGTAAACACAGGTCCCACCAGCTCCATGGGA
GTGACCACCTCAGTCCAGGGAAAAACAGCTGAAGACCAAGATGGTGAGCTCTGATTGCTTCAGTTGGTCATCAACTATT
TTCCCTTGACTGCTGTCTGGGATGGCCTGCTATCTTGATGATAGATTGTGAATATCAGGAGGCAGGGATCACTGTGG
ACCATCTTAGCAGTTGACCTAACACATCTTCTTTTCAATATCTAAGAACTTTTGCCACTGTGACTAATGGTCCTAATA
TTAAGCTGTTGTTATATTTATCATATATCTATGGCTACATGTTATATTATGCTGTGGTTGCGTTTCTGTTTATTTA
CAGTTGCTTTTACAAATATTTGCTGTAACATTTGACTTCTAAGGTTTAGATGCCATTTAAGAACTGAGATGGATAGCT
TTTAAAGCATCTTTTACTTCTTACCATTTTTTAAAGTATGCAGCTAAATTCGAAGCTTTTGGTCTATATTGTTAATT
GCCATTGCTGTAAATCTTAAATGAATGAATAAAATGTTTCATTTTACAAAAA

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSPCDEKKQNDVIAECSNRRLQEVPTVGKYV
TELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHONGNPGIQSNGLNITDGAFNLNLKNLREL
LLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKVCEKTNIEDG
VFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFSLNTQIKYISEEDFKGLINLTLLDLSGNCPRC
FNAPFPCVPCDGGASINIDRFAFQNLTLRLYNLSSTSLRKINAAWFKNMPHLKVLDLEFNVLVG
EIVSGAFLTMLPRLEILDLSFNIIKGSYPQHINISRNFSKLLSLRALHLRGYVFQELREDDFQPL
MQLPNLSTINLGINFIKQIDFKLFQNFNLEIIYLSNRISPLVKDTRQSYANSSSFQRHIRKRR
STDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFIGPNQFENLPDIACLNLSANSNAQV
LSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDEVLDDLSYNSHYFRIAGVTHHLEFIQNFTN
LKVNLNSHNNIYTLTDKYNLESKSLVELVFSGNRLDILWNDDDNRYISIFKGLKNLTRDLNLR
LKHIPNEAFLNLPASLTELHINDNMLKFFNWTLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLR
TLLLSHNRISHLPSGFLSEVSSLKHLDDLSSNLLKTINKSALETKTTTKLSMLELHGPNPFECTCDI
GDFRRWMDEHLNVKIPRLVDVICASPGDQRGKSIVSLELTTCVSDVTAVILFFFTFFITTMVMLA
ALAHHLFYWDVWFIYNVCLAKVKGYRSLSTSQTIFYDAYISYDTKDASVTDWVINELRYHLEESRD
KNVLLCLEERDWDPLGLAIIDNLMQSINQSKKTVFVLTKKYAKSWNFKTAFYLLALQRLMDENMDVI
IFILLEPVLQHSQYLRLRQRICKSSILQWPDNPKAEGLEFWQTLRNVVLTENDSRYNMYVDSIKQ
Y

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGA
CCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCC
AGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAG
GGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCATC
AGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCCCT
GTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGG
CACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGC
ACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCCTGTCTC
CGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCCTGCA
GCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCCT
CGCTACGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAAT
ATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAG
GATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCC
CAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGC
AGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGG
ACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTG
CAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCC
CGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTT
CCTTCCTGGAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCA
GGCTGGACTGAGCCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCC
AGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCCT
TCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCT
GGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAG
GTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGG
GCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAAATGAAACGTGAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTT
GGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRACSTY
RTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRRCRCPAGWR
GDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPGGGPPRVAPNPTGVDSA
MKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFL
EEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGC
TAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGC
ATCAGAGCAGCCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCCGCTGTGAGGGGCTTCG
CGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGCAGTGGATGAGCA
ACCCAACGGGGGGCCCGGGGAGGGGAAGTGGCCCCGAGGGAGAGGAACCCCAAAGCCACATCTGTA
GCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCA
CCCGGAGGAGCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGA
GGAGAAGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGG
CTTCTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCCTCACCA
CCTGCGACGGGGACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGC
CCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCT
TCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGC
CTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGC
AGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTG
TTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGGCCCCCAGGG
TGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGACAGAGGCTGCAGTCCAGG
GTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGC
ACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCA
TCGACTCCCTGAGCGAGCAGATTTCTTCCTGGAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAA
GACTCGTGAACTGCCCAGCGCTCCAGGCTGGACTGAGCCCCCTACGCCGCCCTGCAGCCCCCATGC
CCCTGCCCAACATGCTGGGGGTCCAGAAGCCAACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGG
CCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGG
GATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCC
AGGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGC
CAGGCAGCCCGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAA
ATGAAACGTG

FIGURE 215

MRGSEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRACSTY
RTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRRCRCPAGWR
GDTQCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSadgtLCVPKGGPPRVAPNPTGVDSA
MKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFL
EEQLGSCSCKDS

Signal sequence:

1-19

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCAGGT
GGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGGTCCATC
TCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCATCAGCAGCCC
CCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCCCTGTCCGGGG
GATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCC
ATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCA
CGCCTACCGGCCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCTGTCTCCGAGTCGT
TCGTGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGC
GTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGC
CGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGG
GGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTG
CGTCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTA
CACTCTGTGTGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCA
ATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGT
GCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC
TCCTGGTGCCTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCCTG
GAGGAGCAGCTGGGGTCTCTGCTCCTGCAAGAAAGACT**CGTGA**CTGCCAGCGCCCCAGGCTGGAC
TGAGCCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAGAAGCCA
CCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGG
AGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCTGGCTACCC
CCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTACGAGC
TCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGGCCTCAGT
GGGGGCTGCTGCCTGACCCCCAGCACATAAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRACSTY
RTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRRCRCPAGWR
GDTQCSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSADGTL CVPKGGPPRVAPNPTGVDSA
MKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFL
EEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 218

GGTTGCCACAGCTGGTTTAGGGCCCCGACCACTGGGGCCCCCTTGTCAGGAGGAGACAGCCTCCCGGCCCGGGGAGGAC
AAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCCTGGGACGGTCCGTTTCCTGCCGTGAGCTGCCGGCCGAGTTGG
GTCTCCGTGTTTTAGGGCCGGCTCCCCCTTCTGGTCTCCCTTCTCCCGCTGGGCCGGTTTATCGGGAGGAGATTGTCT
TCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGATTTCAAAGCTGGG
CTCAGCCTCTGTTTCTTCTCTCGTGAATCGCAAAACCCATTTTGGAGCAGGAATCCAATCATGTCTGTGATGGTGG
TGAGAAAGAAGGTGACACGGAAATGGGAGAACTCCCAGGCAGGAACACCTTTTGCTGTGATGGCCGCGTCATGATGG
CCCCGCAAAAGGGCATTCTTACCTGACCCTTTTCTCATCTCTGGGGACATGTACACTCTTCTTCGCCCTTTGAGTGCC
GCTACCTGGCTGTTTCACTGTCTCCTGCCATCCCTGTATTTGCTGCCATGCTCTTCCTTTTCTCCATGGCTACACTGT
TGAGGACCAGCTTCAGTGACCCTGGAGTGATTCTCTGGGCGCTACCAGATGAAGCAGCTTTCATAGAAATGGAGATAG
AAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGCTCGTATCAAGAAATTTCCAGATAAACAACCAGATTG
TGAAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCCGGGCCTCCCATTCGAGCATCTGTGACAACTGTG
TGGAGCGCTTCGACCATCACTGCCCCTGGGTGGGGAAATTGTGTTGGAAAGAGGAACCTACCGTACTTCTACCTCTTCA
TCCTTTCTCTCTCCCTCTCACAACTATGTCTTTCGCTTCAACATCGTCTATGTGGCCCTCAAATCTTTGAAAATTG
GCTTCTTGGAGACATTGAAAGAACTCCTGGAACGTGTTCTAGAAGTCTCATTGCTTCTTACACTCTGGTCCGTCG
TGGGACTGACTGGATTTTCTACTTTCTCTCGTGGCTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAG
GGAAGAATCGCGTCCAGAATCCCTACAGCCATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGTGGCCCCCTTGC
CCCCCAGTGTGCTGGATCGAAGGGGTATTTTGGCACTGGAGGAAAGTGGAAGTCGACCTCCAGTACTCAAGAGACCA
GTAGCAGCCTCTTGCCACAGAGCCAGCCCCACAGAACACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTC
CCGAAGAGATGCCACCTCCAGAGCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCTATCTATGGAA
GAGACTTTTGTGTTGTGTTAATTAGGGCTATGAGAGATTTTCAAGGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAG
CTGTCCCTTTTAACTGTTTTTCTTTGGTCTTTAGTCAACCCAGTTGCACACTGGCATTCTTCTGCTGAAGCTTTTTTA
AATTTCTGAACTCAAGGCAGTGGCAGAAGATGTGAGTCACTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGC
ACTGGTTCTCCATGGCCTCAGCCACAGGGTCCCTTGGACCCCTCTCTTCCCTCCAGATCCAGCCCTCCTGCTTGG
GGTCACTGGTCTCATTCTGGGGCTAAAAGTTTTTGGAGACTGGCTCAAATCCTCCCAAGCTGCTGCACGTGCTGAGTCC
AGAGGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGGTCAGAAGATTCTCCTGGCCACCAAGTGCCAGCATTGCCACAAATCCTTTTAGGAATGGGACAGGTACCTTCC
ACTTGTTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNTTGTGTTTTCTTTTGAAGTCTGCTCCCATTAGGAGCAGGAA
TGGCAGTAATAAAAGTCTGCACTTTGGTCACTTTCTTTTCTCAGAGGAAGCCCGAGTGCTCACTTAAACACTATCCCC
TCAGACTCCCTGTGTGAGGCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAGGCTCTCCTCTCC
TCTCCTCTCCCCGATGTACCCTCAAAAAAAAAAAAAATGCTAACAGTTCTTCCATTAAAGCTCGGCTGAGTGAGGGA
AAGCCCAGCACTGCTGCCCTCTCGGGTAACTCACCTAAGGCCTCGGCCACCTCTGGCTATGGTAACCACTGAGG
GCTTCTTCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCCAGAGCCACTTACCCTGGGGGTGGGCTGTGG
CCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTTCAAGGAAGAAGATTTATGTATTATATGTGGCTATATTTCTAG
AGCAGCTGTGTTTTCTCTTTCTAAGCCAGGGTCTGTCTGGATGACTTATGCGGTGGGGGAGTGTAACCGGAACCTT
TTCATCTATTTGAAGCGATTAACTGTGTCTAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTFLILGTCTLFFAFECRYLAVQLS
PAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPPRIKNFQ
INNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGRNYRYFYLFILSLSLT
IYVFAFNIVYVALKSLKIGFLETCLKETPGTVLEVLCFFTLWSVVGLTGFHTFLVALNQTNEDI
KGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSRPPSTQETSSSLLPQS
PAPTEHLNSNEMPEDSSTPEEMPPEPPEPPQEAAEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTATTAT
AGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCTCCCACA
GAGCNCTTCGACCATCACTGCCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTACCGCTANTT
CTACCTCTTCATCCTTTNTCTCTCCCNCTCACAATCTATGTCTTCGCCTTCAACATCGT

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTAAATCTCCTTGACACAAGCTTGAGAGCAACACAATCT
ATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGAAAAAA
ATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCACGGGGCTGGC
TGCTCTGTGTCTCTTCCAAGGAGTGGCCGTGCGCAGCGGAGATGCCACCTTCCCCAAAGCTATGG
ACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATTGACAACCGGGTCACC
CGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGACAAGTGGTGCCTGGATCC
TCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCGAGATCCAGAACGTGGATGTGT
ATGACGAGGGGCCCTTACACCTGCTCGGTGCAGACAGACAACCACCCAAAGACCTCTAGGGTCCAC
CTCATTGTGCAAGTATCTCCCAAATTTGTAGAGATTTCTTCAGATATCTCCATTAATGAAGGGAA
CAATATTAGCCTCACCTGCATAGCAACTGGTAGACCAGAGCCTACGGTTACTTGGAGACACATCT
CTCCCAAAGCGGTTGGCTTTGTGAGTGAAGACGAATACTTGGAAATTCAGGGCATCACCCGGGAG
CAGTCAGGGGACTACGAGTGCAGTGCCTCCAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAA
GGTCACCGTGAACCTATCCACCATACATTTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGGACAAA
AGGGGACACTGCAGTGTGAAGCCTCAGCAGTCCCCTCAGCAGAATTCAGTGGTACAAGGATGAC
AAAAGACTGATTGAAGGAAAGAAAGGGGTGAAAGTGGAAAACAGACCTTCTCTCAAACTCAT
CTTCTTCAATGTCTCTGAACATGACTATGGGAACTACACTTGCGTGGCCTCCAACAAGCTGGGCC
ACACCAATGCCAGCATCATGCTATTTGGTCCAGGCGCCGTGAGCGAGGTGAGCAACGGCACGTGCG
AGGAGGGCAGGCTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGACCTGCTTCTCAAATTTTTGATG
TGAGTGCCACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCAACACAACAGCAATGGC
AACACCGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACA
GAAATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAAA
ATTGCCTTGACAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGCACAC
CCGGCTTGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAAGGGCTCA
GCCTCTCTGCCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCAAATTCAATCA
GTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGGGCACTTTGGTAGA
CTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAAAAAA

FIGURE 222

MKTIQPKMHNSISWAI FTGLAALCLFQGV PVRSGDATFPKAMDNVTVRQGESATLRCTIDNRVTR
VAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNV DVYDEGPYTCSVQTDNHPKTSRVHL
IVQVSPKIVEISSDISINEGNNISLT CIATGRPEPTVTWRHISPKAVGFVSEDEYLEIQGITREQ
SGDYEC SASNDVAAPV VRRVKVTVNYPPYI SEAKGTGVPVGQKGT LQCEASAVPSAEFQWKDDK
RLIEGKKGVKVENRPFLSKLIF FNVSEHDYGN YTCVASNKLGH TNASIMLFGPGAVSEVSNGTSR
RAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28.

FIGURE 223

GAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCACG
GGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCCCAA
AGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATTGACAACC
GGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGACAAGTGGTGC
CTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCGAGATCCAGAACGT
GGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAACCACCCAAAGACCTCTA
GGGTCCACCTCATTGTGCAAGTATCTCCCAAATTTGTAGAGATTTCTTCAGATATCTCCATTAAT
GAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGACCAGAG

FIGURE 224

ATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCGGGGCCCGGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATC
ACCGCCTGGCCCGACTCCACCATGAACGTGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGGG
ACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCTGCACTGCTT
CTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCACAGCACCTGCCTTACAGAGGCCTGCATT
CGAGTGGCTGGAAAAATCTGGAGTCCCTGGACCGAGGGGTGAGCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGG
GGCTGGATTCCGAGGAACCCCTGCCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAACCAG
GCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCAGAAGACACAGCGCTTCTAC
CTATCTTGCCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCCAGCCACTGAGAGACCTCATTGAGAAGATTGGTGGT
TGGAACATTACGGGGCCCTGGGACCAGGACAACCTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACC
CCATTCTTACCAGTCTACATCAGTGCCGACTCTAAGAGTTCCACAGCAATGTTATCCAGGTGGACCACTCTGGGCTC
TTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTGCTCACTGCCTATCTGGATTACATGGAG
GAACTGGGGATGCTGCTGGGTGGGCGGCCACCTCCACGAGGGAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAG
CTGGCCCAACATCACAGTGCCCAAGGACCAGCGGCGCGACGAGGAGAAGATCTACCACAAGATGAGCATTTCGGAGCTG
CAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTCTGTCTTTCTTGCTGTCACCATTTGGAGTTGAGTGACTCTGAG
CCTGTGGTGGTGTATGGGATGGATTATTTGCAGCAGGTGTCAGAGCTCATCAACCGCACGGAACCAAGCATCCTGAAC
AATTACCTGATCTGGAACCTGGTGCAAAAGACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTG
CTGGAGACCTCTATGGCACTAAGAAGTCTGTGTGCGGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTT
GGCTTTGCTTTGGGGTCACTCTTCGTGAAGGCCACGTTTGACCGGCAAGCAAAGAAATTGCAGAGGGGATGATCAGC
GAAATCCGGACCGCATTGTAGGAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAG
AAAGCAGATGCCATCTATGATATGATTGGTTTCCAGACTTTATCCTGGAGCCCAAAGAGCTGGATGATGTTTATGAC
GGGTACGAAATTTCTGAAGATTCTTCTTCCAAAACATGTTGAATTTGTACAACCTTCTCTGCCAAGGTTATGGCTGAC
CAGCTCCGCAAGCCTCCCAGCCGAGACCAGTGGAGCATGACCCCCAGACAGTGAATGCCTACTACCTTCCAACCTAAG
AATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCCTTCTATGCCCGCAACCAACCCCAAGGCCCTGAAC'TTCGGT
GGCATCGGTGTGGTGTATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAAC
CTGCGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCAACAGGCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGCGCCGCCAGACGCTGGGGGAGAACATTACTGACAACGGGGGGCTGAAGGCT
GCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCAC
CAGCTCTTCTTCTGTTGGATTGTGCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACAGGGGGCTGGTGACC
GACCCCAACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAAC'TCCCGTGACTTCTGCGGCACTTCGGCTGC
CCTGTGCGCTCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACCTGGATCAGGGGAGAAATGGCCAGCTGT
CACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGT
CTAGTCCCTCCCCCCCCACAGGTGACATGAGTACAGACCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCC
CCTGCCTCCAGCAGAGCCCCACCATTCACTGTGACATCTTTCCGTGTCACCCTGCCTGGAAGAGGTCTGGGTGGGA
GGCCAGTTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSHSTC
LTEACIRVAGKILESIDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQAILKHL
ENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLEKIGGWNITGPWDQDNFMEVLKAVA
GTYRATPFFTVYISADSKSSNSNVIQVDQSGFLFLPSRDYYLNRTANEKVLTAAYLDYMEELGMLLG
GRPTSTREMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAPSMDWLEFLSFLLSPL
ELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSLDLRRFESAQEKLETLTLYGT
KKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMISEIRTAFFEEALGQLVWMDEKT
RQAAKEKADAIYDMIGFPDFILEPKELDDVDGYEISEDSFFQNMLNLYNFSAKVMADQLRKPPS
RDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNHPKALNFGGIGVVMGHELTHAFDDQGRE
YDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQVNGERLNGRQTLGENITDNGGLKAAYNAYKA
WLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCSVRTPESSHEGLVTDPHSPARFRVLGTLSNSRDF
LRHFGCPVGSPMNPGLCEVW

Type II Transmembrane domain:

amino acids 32-57

FIGURE 226

GCCCGGCCCTCCGCCCTCCGCACTCCCGCCTCCCTCCCTCCGCCCCGCTCCCGCGCCCTCCCTCCCTCCCTCCCCAG
CTGTCCCGTTCCGCTCATGCCAGCCTCCCGGCCCGCCCGGCCCGCTGCTGCTCCTCGGGCTGCTGCTCGGCTC
CCGGCCGGCCCGCGGCCCGGCCAGAGCCCCCGTGTGCTGCCATCCGTTCTGAGAAGGAGCCGCTGCCCGTTCCGGG
AGCGGCAGGTAGGTGGCGCCCGGGGAGGCGCGGGCGGGGAGTCCGGCTCGGGGCGAGTCAGCGCCAGCCCGGAGGG
GGCGCGGGGCGCAGGTGGCTCGGCGCGGCGGGCGGCCCGAGGGTGGGCGGGGGCAGAAGGGCGCGGTGCTGGGACC
CGGGACCCCGCGGGCAGCCCCCGGGGCGGCACACGGCGGAGCTGGGCGAGCGGCTCCAGCCAAGCCCGTCCCCGAGG
CTGCACCTTCGGCGGGAAGGTCTATGCCTTGGACGAGACGTGGCACCCCGGACCTAGGGGAGCCATTCCGGGTGATGCG
CTGCGTGTCTGCGCCTGCGAGGCGCAGTGGGGTTCGCCGTACCAGGGGCCCTGGCAGGGTCAGCTGCAAGAATCAAA
ACCAGAGTGCCCAACCCCGGCCCTGTGGGCGAGCCGCGCCAGCTGCCGGGACACTGCTGCCAGACCTGCCCCAGGACTT
CGTGGCGCTGCTGACAGGGCCGAGGTGCGCAGGCGGTGGCAGGAGCCGAGTCTCGCTGCTGCGCTCTAGCCTCCGCTT
CTCTATCTCCTACAGGCGGCTGGACCGCCCTACCAGGATCCGCTTCTCAGACTCCAATGGCAGTGTCTCTGTTGAGCA
CCCTGCAGGTGCTGCCAACCTGACAGTCCAGGAGATGGACTGGCTGGTGTGGGGGAGCTGCAGATGGCCCTGGAGTG
GGCAGAACAGCTGCATGTGGCACTTGTGACACTCACTACCCCTCAGGGGAGGTCTGGGGGCTCTCATCCGCGACCG
GGCCCTGTCCCCAGAGACCTCAGTGCCATCCTGACTCTAGAAGGCCCCACCAGCAGGGCGTAGGGGGCATCACCTT
GCTCACTCTCAGTGACACAGAGGACTCCTTGCAATTTTTGCTGCTCTTCCGAGGCCTTGCAAGACTAACCCAGGTTC
CTTGAGGCTCCAGATTCTACACCAGGGGAGCTACTGCGAGAACCTCAGGCCAATGTCTCAGCCAGGAACCCAGGCTT
TGCTGAGGTGCTGCCAACCTGACAGTCCAGGAGATGGACTGGCTGGTGTGGGGGAGCTGCAGATGGCCCTGGAGTG
GGCAGGCAGGCCAGGGCTGCGCATCAGTGACACATTGCTGCCAGGAAGAGCTGCGACGTCTGCAAGTGTCTTTG
TGGGGCTAATGCCCTGATCCAGTCCAAACGGGTGCTGCCGGCTCAGCCAGCCTCACTCTGCTAGGAAATGGCNCCT
GATCCTCCAGGTGCAATTGGTAGGACAACCAAGTGGTGGTGGCCATGACACTGGAACCAAGCCTCAGCGGAGGGA
TCAGCCCACTGTCTGTGCCACATGGCTGGCTATCCTCCCTGCCCGCCAGGCCGTGGGTATCTGCCCTGGGCTGGGG
TGCCCGAGGGGCTCATATGCTGCTGCAGAATGAGCTCTTCTGAACGTGGGCACCAAGGACTTCCAGACGGAGAGCT
TCGGGGGCAACGTGGCTGCCCTGCCCTACTGTGGGGCATAGCGCCCGCTGCCCGTGGCCCTAGCAGGAGCCCTGGT
GCTACCCCTGTGAAGAGCCAAGCAGCAGGGCAGCCTGGCTTTCCTTGATACCCACTGTCACTGTCACTATGAAGT
GCTGCTGGCTGGGCTTGGTGGCTCAGAACAAGGCACTGTCACTGCCACCTCCTTGGGCCCTCCTGGAACGCCAGGGCC
TCGGCGGCTGCTGAAGGATTTCTATGGCTCAGAGGCCAGGGTGTGGTGAAGGACCTGGAGCCGGAAGTGTGCGGCA
CCTGGCAAAAGGCATGGCTTCCCTGATGATCACCACCAAGGTAGCCCCAGAGGGGAGCTCCGAGGGCAGCCTCTCTC
CCAGGTGCACATAGCCAACCAATGTGAGGTGGCGGACTGCGCCTGGAGGCGGCCGGGGCCGAGGGGGTGGGGCGCT
GGGGGCTCCGGATACAGCCTCTGCTGCGCCGCTGTGGTGCCTGGTCTCCCGGCCCTAGCGCCCGCCAAACCTGGTGG
TCCTGGGCGGCCCCGAGACCCCAACACATGCTTCTTCGAGGGGAGCAGCGCCCCCACGGGGCTCGCTGGGCGCCCAA
CTACGACCCGCTCTGCTCACTCTGCACCTGCCAGAGACGAACGGTGATCTGTGACCCGGTGGTGTGCCACCGCCAG
CTGCCCCACACCCGGTGAGGCTCCCGACCACTGCTGCCCTGTTTGGCCTGGCTGCTATTTGATGGTGACCGGAGCTG
GCGGGCAGCGGGTACGCGGTGGCACCCCGTTGTGCCCCCTTTGGCTTAATTAAGTGTGCTGTCTGCACCTGCAAGCA
GGGGGGCACTGGAGAGGTGCACTGTGAGAAGGTGCACTGTCCCCGGCTGGCCTGTGCCCAGCCTGTGCGTGTCAACCC
CACCAGTGTGCAAAACAGTGTCCAGGTGAGGCCCAACCCAGCTGGGGGACCCCATGCAGGCTGATGGGCCCCGGGG
CTGCCGTTTTGTGTTGGCAGTGGTTCAGAGAGTCAAGCTGGCACCCCTCAGTGCCCCGTTTGGAGAGATGAGCTG
TATCACTGCAGATGTGGGGTAAGTGGGGAGCAGAGGCTTGTGTGAGGTGGTACTGGGAGCCTGGTCTGGAGTAGGG
AGACCTTCCAGGGAGGTCCCTGAAGAAGCTGAAGGTCACTGTGTCCAGTGCCCTTGGGGGACACTCAGTGTCTGCT
CTGTCTTGTACCAGGCAGGGGTGCCCTCACTGTGAGCGGGATGACTGTTCACTGCCACTGTCTGTGGCTCGGGGAAGG
AGAGTCGATGCTGTTCCCGCTGCACGGGCCACCGCGCGCTAAGTGAAGGAGTCCAGGGTCAGCAGCTGTGAGTGGAG
GGCTCACCTGCCCTGTGGGACTCCTGATCAGGGAAGGGAGCACTCACTGTGTGCAGGAACAGTGCAGCCTGCCTCACA
GTGCCATTCCAATCCACCCTCAGCAACCTGGTGAATTTGTTATTTATGACCTTTCTTTACAAATGAGATTTCTGA
AGCTCAGAGAAATTAAGCAACGAGATGAAGGTCAACCCAGCTGTGTGCACTGACCTGTTAGAAAATACTGGCCTTTCT
GGGACCAAGGCAGGGATGCTTGGCCCTGCCCTCTATGCCTCTCTGTGCCTCTCCACTCCCTCTCCCTCCTCCAACAT
TCCCTCCCTTCTGTCTCCAGCAGCCCCAGAGACCAGAAGTATCCAGAGCTGGAGAAAGAAGCCGAAGGCTCTTAGGG
AGCAGCCAGAGGGCCAAGTGACCAAGAGGATGGGGCTGAGCTGGGGAAGGGGTGGCATCGAGGACCTTCTTGCAATC
TCCTGTGGGAAGCCAGTGCCCTTGTCTCTGTCTGCTCTACTCCACCCCACTACCTCTGGGAACCAAGCTC
CACAAGGGGGAGAGGCAGCTGGGCCAGACCGAGGTACAGCACTCCAAGTCTGCCCCTGCCACCCCTCGGCCTCTGT
CTGGAAGCCCCACCCCTTTCTTCTGTACATAATGTCACTGGCTTGTGGGATTTTAAATTTATCTTCACTCAGCACC
AAGGGCCCCGGCACTCCACTCCTGCTGCCCTGAGCTGAGCAGAGTCATTATTGGAGAGTTTGTATTTATTAAC
ATTTCTTTTTTCACTCTTGGGCATGAGGTTGGCTCTTTGTGGCCAGGAACCTGAGTGGGGCTGGTGGAGAGGGGCGN
GAGAGTAGGAGGTGAGAGAGAGGAGCTCTGACACTTGGGGAGCTGAAAGAGACCTGGAGAGGCAGAGAGCGTGGC
NNTTGGCTGGCATNCTGGGTTCCGAGAGGGGTGGGGATGGTTCTTGAGATGGTCTAGAGACTCAAGAATTTAGGG
AAGTAGAAGCAGGATTTTGAAGTCAAGTTTCCACATCGCTGGCCTGTTGCTGACTTCATGTTGAAGTTGCT
CCAGAGAGAGAATCAAAGGTGTACCAGCCCTCTCTCCCTCCTTCCCTTCCCTTCTTCTTCCCTCCCTCCCTCC
TCCCTCCCTCCCTCC

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCG
GGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCCGACTCCACC
ATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGGGACAAGACA
GCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCTGCAC
TGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCCACAGCACCTGC
CTTACAGAGGCCTGCATTTCGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCC
CTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTTCGGAGGAACCCCTGCCCGATGGGC
GTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTT
GAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCAGAAGACACAGCGCTTCTACCTATCTTG
CCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCCAGCCACTGAGAGACCTCATTGAGAAGATTG
GTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACCTTTATGGAGGTGTTGAAGGCAGTAGCA
GGGACCTACAGGGCCACCCCATTTCTTCACCGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAG
CAATGTTATCCAGGTGGACCAGTCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAA
CTGCCAATGAGAAAGTAAGGAACATCTTCCGAACCCCATCCCTACCCCTGGCTGAGCTGGGCTG
ATCCCTGTTGACTTTTCCCTTTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTACCTA
GTGAACAACTGCCCCTCCTTTCTTTCTTTCTTTCTTCCCTCCCTCCCTTTCTTCCCCTTTT
CCTTCCTTCCTTCCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGT
GGGGGAATTCGAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTCGATGGGTTC
ATGGACCTAGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTCAGGAGAGACTTATG
GAGCCAGCAAAGTCTTCCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

FIGURE 228

ATGCCTACTACCTTCCAAC TAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCCTTC
TATGCCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCATGAGTT
GACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGCCCTGGTGGC
AGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAGTACAATCAATAC
CAGGTCAATGGGGAGAGGCTCAACGGCCCGCCAGACGCTGGGGGAGAACATTGCTGACAACGGGGG
GCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGC
CAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTTGCCCAGGTGTGGTGCTCGGTC
CGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCCCCACAGCCCTGCCCCGCTTCCGCGT
GCTGGGCACTCTCTCCAAC TCCCGTGACTTCCTGCGGCACTTCGGCTGCCCTGTGCGCTCCCCCA
TGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACCTGGATCAGGGGAGAAATGGCCAGCTGTAC
CAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGCTCTTGGGTGAGGAAGCAAATGCAAGC
TGGGCTGGGTCTAGTCCCTCCCCCCCACAGGTGACATGAGTACAGACCCTCCTCAATCACCACAT
TGTGCCTCTGCTTTGGGGGTGCCCCTGCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTT
CCGTGTCAACCCTGCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCT
CTTCTGTCCCCAGGCTCACTCAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACC
CACAGGCCTGGGTGGTGTACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTG
GACTCAGCTCTGTCTGGCTCACCTCACGGGCTACCCCCACCTCACCTGTGCTCCTTGTGCCAC
TGCTCCCAGTGCTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAG
CCTCCTGCTGCCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTT
CCTGTGTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGA
TAGAGCAGGGAAAAGGAAGAACAGAGTTTATTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCTT
GGCCCTTATAGGACC

FIGURE 229

CCCACGCGTCCGAGCCGCCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGAGGA
GGGGAGGCAAAAACACCGAAAAACAAAAAGAGAGAAACAACACCCCAACAACCTGGGGTGGGGGAA
GAAAGAAAGAAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAATCCTGTGGCGCGCCCGCTGGTTCCCGGGAAGACTCGCCAGCACCAGGGGGTGGGGGAGTGCG
AGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATGCTGTTGGTGCAGGGT
GCTTGTGTGCTCGAACCAGTGGCTGGCGGGCGGTGCTCCTCAGCCTGTGCTGCCTGCTACCCTCCTG
CCTCCCGGCTGGACAGAGTGTGGACTTCCCCTGGGCGGCCGTGGACAACATGATGGTCAGAAAAG
GGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCTTCAAAGGGTGCCTGGCTGAACCGG
TCAAGTATTATTTTTGCGGGAGGTGATAAGTGGTCAGTGGATCCTCGAGTTTCAATTTCAACATT
GAATAAAAGGGACTACAGCCTCCAGATACAGAATGTAGATGTGACAGATGATGGCCCATACACGT
GTTCTGTTTCAGACTCAACATACACCCAGAACAATGCAGGTGCATCTAACTGTGCAAGTTCCTCCT
AAGATATATGACATCTCAAATGATATGACCGTCAATGAAGGAACCAACGTCACTCTTACTTGTTT
GGCCACTGGGAAACCAGAGCCTTCCATTTCTTGGCGACACATCTCCCATCAGCAAAACCATTTG
AAAATGGACAATATTTGGACATTTATGGAATTACAAGGGACCAGGCTGGGGAATATGAATGCAGT
GCGGAAAATGCTGTGTCATTTCCAGATGTGAGGAAAGTAAAGTTGTTGTCAACTTTGCTCCTAC
TATTTCAGGAAATTAAATCTGGCACCGTGACCCCGGACGCAGTGGCCTGATAAGATGTGAAGGTG
CAGGTGTGCCGCTCCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAA
GGAATTATTATTCAAAATTTTAGCACAAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCA
CTTCGGCAATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTA
ACCCTCCAAGTACAGCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTTCTCCTGCTGGTAC
CTTGTGTTGACACTGTCTCTTTTACCAGCATATTCTACCTGAAGAATGCCATTCTACAATAAAT
TCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCTGGTA
CAGTTTGTAAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCCTTCTGTA
GAATTGCTCATTATGTAAATACTTTAATTCTACTCTTTTTTGTATTAGCTACATTACCTTGTAAG
CAGTACACATTGTCTTTTAAAAAGCGTGAAAGCTCTGAAATTACTTTTAGAGGATATTAATTG
TGATTTTCATGTTTGTAAATCTACAACCTTTTCAAAGCATTCACTCATGGTCTGCTAGGTTGCAGGC
TGATGTTTACAAAAACGAATATTGCAGTGAATATGTGATTCTTTAAGGCTGCAATACAAGCATTC
AGTTCCCTGTTTCAATAAGAGTCAATCCACATTTACAAAGATGCATTTTTTTTCTTTTTTGATAAA
AAAGCAATAATATTGCCTTCAGATTATTTCTTCAAATATAACACATATCTAGATTTTTTCTGCT
TGCATGATATTTCAGGTTTCAGGAATGAGCCTTGTAATATAACTGGCTGTGCAGCTCTGCTTCTCT
TTCTGTAAAGTTTCAGCATGGGTGTGCCTTCATACAATAATATTTTTTCTTTTGTCTCCAACATA
ATAAAATGTTTTGCTAAATCTTACAATTTGAAAGTAAAAATAAACAGAGTGATCAAGTTAAACC
ATACACTATCTCTAAGTAACGAAGGAGCTATTGGACTGTAAAAATCTCTTCCCTGCACTGACAATG
GGGTTTGAGAATTTTGCCCCACACTAACTCAGTTCTTGTGATGAGAGACAATTTAATAACAGTAT
AGTAAATATACCATATGATTTCTTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTC
CCTTTAAATGACAGCACAGTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTTTCTTTTTCAC
TAGTCCAAGCCAAAAATTTTAAGATGATTTGTGAGAAAGGACAAAGTCCATACCTAATATT
ACAAGAGTTGGTAAGCGCTCATCATTAATTTTATTTTGTGGCAGGTATTATGACAGTCGACCTGG
AGGGTATGGATATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTTATGACC
GCTACTCAGGAGGAAATTACAGAGACAATTATGACAACTGAAATGAGACATGCACATAATATAGA
TACACAAGGAATAATTTCTGATCCAGGATCGTCTTCCAAATGGCTGTATTTATAAAGGTTTTTG
GAGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTTAAATTGACCTGCCAAG
GTAGCTGAAGACCTTTTAGACAGTTCCATCTTTTTTTTTTAAATTTTTTCTGCCTATTTAAGACA
AATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLA AVL LSLCCLLP SCLPAGQSVDFPWA AVDNMMVRK GDTAVLR CYLEDGAS
KGAWLNRSSII FAGGDKWSVDPRVSISTLNKRDYSLQIQNV DVTDDGPYTCSVQTQHTPRTMQVH
LTVQVPPKIIDISNDMTVNEGTVNLTLCLATGKPEPSISWRHISPSAKPFENGQYLDIYGITRDQ
AGEYECSAENAVSFPDVRKV KVVVNFAPT IQEIKSGTVTPGRSGLIRCEGAGVPPPAFEWYKGEK
KLFNGQQGII IQNFSTRSILTVTNVTQEHFGNYTCVAANKL GTTNASLPLNPPSTAQYGITGSAD
VLFSCWYLVLTLS SFTSIFYLKNAILO

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTCTGCTGGCTC
CAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTCTCAAGGAGGACGAGCT
CTTGAGTGAGACCCAACAAGCTGCTTTTACCAAATTGCAATGGAGCCTTTCGAAATCAATGTTT
CAAAGCCCAAGAGGAGAAATGGGGTGAACCTTCTCCCTAGCTGTGGTGGTCATCTACCTGATCCTG
CTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGCGCGGCTCCGGGTCTT
GGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGTCCTTCTCCTTGCTGCAGT
CAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTGCAAGTCCTGCAGGCCCAACTC
ACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGTAGACAACTTCACTCAGAACCAGG
GATGTTTCAGAATCAAAGGTGAACAAGGCGCCCCAGGTCTTCAAGGTCACAAGGGGGCCATGGGCA
TGCCTGGTGCCCCCTGGCCCCGCGGGACCACCTGCTGAGAAGGGAGCCAAGGGGGCTATGGGACGA
GATGGAGCAACAGGCCCCCTCGGGACCCCCAAGGCCACCAGGAGTCAAGGGAGAGGCGGGCCTCCA
AGGACCCCAGGGTGCTCCAGGGAAGCAAGGAGCCACTGGCACCCCAGGACCCCCAAGGAGAGAAGG
GCAGCAAAGGCGATGGGGGTCTCATTGGCCCCAAAGGGGAAACTGGAATAAGGGAGAGAAAAGGA
GACCTGGGTCTCCAGGAAGCAAAGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCTATGGGGCC
TCCTGGAGCCCAGGGGAGTAAAGGTGACTTCGGGAGGCCAGGCCACCAGGTTTGGCTGGTTTTT
CTGGAGCTAAAGGAGATCAAGGACAACCTGGACTGCAGGGTGTTCGGGGCCCTCCTGGTGCAGTG
GGACACCCAGGTGCCAAGGGTGAGCCTGGCAGTGCTGGCTCCCCTGGGCGAGCAGGACTTCCAGG
GAGCCCCGGGAGTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGAC
AGCAAGGAAGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGC
CCAGGGCTGGCAGGTCCCAAGGGAGCCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAA
AGGATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACTCAGTGTCCG
TCAGGATTGTTCGGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGGGACA
ATTTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTTACTCCAA
AGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAATGTTTCAAGTGT
GGGGCACGGAGAGTACCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCATGACTGCAGCCAC
GAGGAGGACGCAGGCGTGGAGTGCAGCGTCTGACCCGAAACCCCTTTCACTTCTCTGCTCCCGAG
GTGTCCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGTTCCCTGGGGACAACCTGAGC
AGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLILLTAGAGLLVV
QVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVRVSHEHL
LQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRDGATGPSGPQ
GPPGVKGEAGLQGPQGAPGKQGATGTPGPQGEKSGKDGGLIGPKGETGTKGEKGDGLPGSKGD
RGMKGDAGVMGPPGAQGSKGDFGRPGPPGLAGFPGAQGDQGPGLQGVPGPPGAVGHPGAKGEPG
SAGSPGRAGLPGSPGSPGATGLKSGKGDGTGLQGQQGRKGESGVPGPAGVKGEQGSPLAGPKGAP
GQAGQKGDQGVKGSSEGEQGVKGEKGERGENSVSVRIVGSSNRGRAEVYYSGTWGTICDDEWQNSD
AIVFCRMLGYSKGRALYKVGAGTGQIWLDNVQCRGTESTLWSCTKNSWGHHDSCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

Clq domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTCAATTTGTAAAGAAGCTCCTTCCAGCACCTCCTCTCTT
CTCCTTTTGCCCAAACCTCACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAGACCAA
AGGAAAGAAGAAAAGGGCCAAAAGCCAAAATGAACTGATGGTACTTGTTTTACCATTTGGGCT
AACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTCTTGCTACAGAAAGATACTAA
AAGATCACAACCTGTCACAACCTTCCGGAAGGAGTAGCTGACCTGACACAGATTGATGTCAATGTC
CAGGATCATTTCTGGGATGGGAAGGGATGTGAGATGATCTGTTACTGCAACTTCAGCGAATTGCT
CTGCTGCCCCAAAAGACGTTTTCTTTGGACCAAAGATCTCTTTCGTGATTCCTTGCAACAATCAAT
GAGAATCTTCATGTATTCTGGAGAACACCATTCTGATTTCCACAAACTGCACTACATCAGTAT
AACTGCATTTCTAGTTTCTATATAGTGCAATAGAGCATAGATTCTATAAATTCTTACTTGTCTAA
GACAAGTAAATCTGTGTTAAACAAGTAGTAATAAAAGTTAATTCAATCTAAAAAAAAAAAAA

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758

<subunit 1 of 1, 98 aa, 1 stop

<MW: 11081, pI: 6.68, NX(S/T): 1

MKLMVLVFTIGLTLLLG VQAMPANRLSCYRKILKDNCHNLPEGVADLTQIDVNVQDHFWDGKGC
EMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGTCTGGAGCGAATTCCAGCCTGCAGGGCTG
ATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCCGCCGTGGTGGTTGGAGGGCGCGCAG
TAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAGATGTGGGAATCTC
CTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCCGCGCTGGCTGTGCGCTGGGGCGCT
GGTGTCTGGCGGGTGGCTTCTTTCTCCTCGGCTTCTCTTCGGGTGGTTTATAAAATCCTCCAATG
AAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTTTGGATGAATTGAAAGCTGAGAAC
ATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGCAGGAACAGAACAAAATTTTCA
GCTTGCAAAGCAAATTCATCCCAGTGGAAAGAATTTGGCCTGGATTCTGTTGAGCTAGCTCATT
ATGATGTCTGTGTCTACCCAAATAAGACTCATCCCACTACATCTCAATAATTAATGAAGAT
GGAAATGAGATTTTCAACACATCATTATTTGAACCACCTCCTCCAGGATATGAAAATGTTTCGGA
TATTGTACCACCTTTCAGTGCTTCTCTCCTCAAGGAATGCCAGAGGGCGATCTAGTGTATGTTA
ACTATGCACGAACTGAAGACTTCTTTAAATTGGAACGGGACATGAAAATCAATTGCTCTGGGAAA
ATTGTAATTGGCAGATATGGGAAAGTTTTTCAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGG
GGCCAAAGGAGTCTTCTACTCCGACCCTGCTGACTACTTTGCTCCTGGGGTGAAGTCCTATC
CAGACGGTTGGAATCTTCTGGAGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCA
GGAGACCTCTCACACCAGGTACCCAGCAAATGAATATGCTTATAGGCGTGGAAATTCAGAGGC
TGTTGGTCTTCCAAGTATTCTGTTCATCCAATTGGATACTATGATGCACAGAAGCTCCTAGAAA
AAATGGGTGGCTCAGCACCACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTT
GGACCTGGCTTTACTGGAACTTTTCTACACAAAAGTCAAGATGCACATCCACTCTACCAATGA
AGTGACGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAGACAGATATGTCA
TTCTGGGAGGTCACCGGGACTCATGGGTGTTTGGTGGTATTGACCCTCAGAGTGAGCAGCTGTT
GTTTCATGAAATTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGGTGGAGACCTAGAAGAACAAT
TTTGTGTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTTCTACTGAGTGGGCAGAGGAGA
ATTCAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTATAGAAGGAAAC
TACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACCTAACAAAAGAGCT
GAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAAGTTGGACTAAAAAAGTCCTT
CCCCAGAGTTCAGTGGCATGCCCAGGATAAGCAAATTTGGATCTGGAAATGATTTTGAGGTGTTT
TTCCAACGACTTGGAATTGCTTCAGGCAGAGCACGGTATACTAAAAATTTGGGAAACAAACAAATT
CAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATATGAGTTGGTGGAAAAGTTTTATGATC
CAATGTTTAAATATCACCTCACTGTGGCCAGGTTTCGAGGAGGGATGGTGTGTTGAGCTAGCCAAT
TCCATAGTGCTCCCTTTTGATTGTGAGATTATGCTGTAGTTTTAAGAAAGTATGCTGACAAAAT
CTACAGTATTTCTATGAAACATCCACAGGAAATGAAGACATACAGTGTATCATTTGATTCACTTT
TTTCTGCAGTAAAGAATTTTACAGAAATTGCTTCCAAGTTCAGTGAGAGACTCCAGGACTTTTGAC
AAAAGCAACCCAATAGTATTAAGAATGATGAATGATCAACTCATGTTTCTGGAAAGAGCATTTAT
TGATCCATTAGGGTTACCAGACAGGCCTTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACA
ACAAGTATGCAGGGGAGTCATTCCCAGGAATTTATGATGCTCTGTTTGATATTGAAAGCAAAGTG
GACCTTCCAAGGCCTGGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTGCAGGC
AGCTGCAGAGACTTTGAGTGAAGTAGCCTAAGAGGATTTTTTAGAGAATCCGTATTGAATTTGTG
TGGTATGTCACTCAGAAAGAATCGTAATGGGTATATTGATAAATTTTAAAATTGGTATATTTGAA
ATAAAGTTGAATATTATATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARRPRWLCAGALVLAGGFLLGFLFGWFIKSSNEATNITPKHNMKAFLDEL
KAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTHPNYISI
INEDGNEIFNTSLFEP PPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFFKLERDMKIN
CSGKIVIARYGKVFRGNKVNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNLPGGGVQGRNILN
LNGAGDPLTPGYPAHEYAYRRGIAEAVGLPSIPVHPIGYYDAQKLEKMGGSAPPDSSWRGSLKV
PYNVGPGFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDYVILGGHRDSWVFGGIDPQS
GAAVVHEIVRSFGLTKKEGWRPRRTILFASWDAEEFLLGSTEWAEENSRLQERGVAYINADSS
IEGNYTLRVDCTPLMYSLVHNLTKELKSPDEGFEGKSLYESWTKKSPSPEFSGMPRISKLGSGND
FEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYETYELVEKFYDPMFKYHLTVAQVRGGMVF
ELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQEMKTYSVSFDLSFSAVKNFTEIASKFSERL
QDFDKSNPIVLRMMNDQLMFLERAFIDPLGLPDRPFYRHVIYAPSSH NKYAGESFPGIYDALFDI
ESKVDPSKAWGEVKRQIYVAAFTVQAAAETLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713

W:\DOCS\BGY\BGY-2104.DOC
030702